Field theory from a bundle point of view

Laurent Claessens

October 13, 2011

http://student.ulb.ac.be/~lclaesse/lectures.pdf

Contents

0.1 What I am

During the firsts days of my thesis, I decided to write down everything I was learning. Some parts of this text were written in 2003 wile others were written yesterday; don't expect a high quality everywhere. This document thus takes the point of view of the learner with some consequences. As far as I can judge my own work:

(i) There are *much* more details in the proofs in this texts that what you can find in other textbooks.

(ii) This is not a text in which you can get a deep understanding of what you are reading.

There are still open questions in the sense that there are points I didn't understand when I wrote. I think that these points are clearly indicated with footnotes or special environment "Problem and misunderstanding". Let me know if you know some answers.

Chapter 1

General differential geometry

1.1 Differentiable manifolds

More precisions in [1], chapter II (§1 and 2) and III. (§1, 2, 3 and 7); [2, 3] can also be useful. An other source for this chapter is [4]. A systematic exposition of manifolds and such can be found in [5].

1.1.1 Definition and examples

A n-dimensional **differentiable manifold** is a set M and a system of charts $\{(\mathcal{U}_\alpha, \varphi_\alpha)\}_{\alpha\in I}$ where each set $\mathcal{U}_\alpha$ is open in $\mathbb{R}^n$ and the maps $\varphi_\alpha\colon \mathcal{U}_\alpha \to M$ are injective and satisfy the three following conditions:

- every $x \in M$ is contained in at least one set $\varphi_\alpha(\mathcal{U}_\alpha)$,
- for any two charts $\varphi_\alpha\colon \mathcal{U}_\alpha \to M$ and $\varphi_\beta\colon \mathcal{U}_\beta \to M$, the set
$$\varphi_\alpha^{-1}(\varphi_\alpha(\mathcal{U}_\alpha) \cap \varphi_\beta(\mathcal{U}_\beta))$$
is an open subset of $\mathcal{U}_\alpha$,
- the map
$$(\varphi_\beta^{-1} \circ \varphi_\alpha)\colon \varphi_\alpha^{-1}(\varphi_\alpha(\mathcal{U}_\alpha) \cap \varphi_\beta(\mathcal{U}_\beta)) \to \mathcal{U}_\beta$$
is differentiable[1] as map from $\mathbb{R}^n$ to $\mathbb{R}^n$.

Each time we say "manifold", we mean "differentiable manifold". We will only consider manifolds with Hausdorff topology (see later for the definition of a topology on a manifold). Any open set of $\mathbb{R}^n$ is a differentiable manifold if we choose the identity map as chart system. Most of surfaces $z = f(x, y)$ in $\mathbb{R}^3$ are manifolds, depending on certain regularity conditions on f.

If M_1 and M_2 are two differentiable manifolds, a map $f\colon M_1 \to M_2$ is **differentiable** if f is continuous and for each two coordinate systems $\varphi_1\colon \mathcal{U}_1 \to M_1$ and $\varphi_2\colon \mathcal{U}_2 \to M_2$, the map $\varphi_2^{-1} \circ f \circ \varphi_1$ is differentiable on its domain. One can show that if $f\colon M_1 \to M_2$ and $g\colon M_2 \to M_3$ are differentiable, then $g \circ f\colon M_1 \to M_3$ is differentiable.

1.1.1.1 Example: the sphere

The sphere S^n is the set
$$S^n = \{(x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \text{ st } \|x\| = 1\}$$
for which we consider the following open set in $\mathbb{R}^n$:
$$\mathcal{U} = \{(u_1, \ldots, u_n) \in \mathbb{R}^n \text{ st } \|u\| < 1\}$$
and the charts $\varphi_i\colon \mathcal{U} \to S$, and $\tilde\varphi_i\colon \mathcal{U} \to S$
$$\varphi_i(u_1, \ldots, u_n) = (u_1, \ldots, u_{i-1}, \sqrt{1 - \|u\|^2}, u_i, \ldots, u_n) \tag{1.1a}$$
$$\tilde\varphi_i(u_1, \ldots, u_n) = (u_1, \ldots, u_{i-1}, -\sqrt{1 - \|u\|^2}, u_i, \ldots, u_n). \tag{1.1b}$$

[1] In the sequel, by "differentiable" we always mean smooth. If this map is differentiable, C^k, analytic,... then the manifold is said to be differentiable, C^k, analytic,...

These map are clearly injective. To see that $\varphi(\mathcal{U}) \cup \tilde{\varphi}(\mathcal{U}) = S$, consider $(x_1, \ldots, x_{n+1}) \in S$. Then at least one of the x_i is non zero. Let us suppose $x_1 \neq 0$, thus $x_1^2 = 1 - (x_2^2 + \ldots + x_{n+1}^2)$ and

$$x_1 = \pm\sqrt{1 - (\ldots)}. \tag{1.2}$$

If we put $u_i = x_{i+1}$, we have $x = \varphi(u)$ or $x = \tilde{\varphi}(u)$ following the sign in relation (1.2). The fact that $\varphi^{-1} \circ \tilde{\varphi}$ and $\tilde{\varphi}^{-1} \circ \varphi$ are differentiable is a “first year in analysis exercise“.

1.1.1.2 Example: projective space

On $\mathbb{R}^{n+1} \backslash \{o\}$, we consider the equivalence relation $v \sim \lambda w$ for all non zero $\lambda \in \mathbb{R}$, and we put

$$\mathbb{R}P^n = \left(\mathbb{R}^{n+1} \backslash \{o\}\right) / \sim .$$

This is the set of all the one dimensional subspaces of $\mathbb{R}^{n+1}$. This is the real **projective space** of dimension n. We set $\mathcal{U} = \mathbb{R}^n$ and

$$\varphi_i(u_1, \ldots, u_n) = \mathrm{Span}\{(u_1, \ldots, u_{i-1}, 1, u_i, \ldots, u_n)\}.$$

One can see that this gives a manifold structure to $\mathbb{R}P^n$. Moreover, the map

$$A\colon S^n \to \mathbb{R}P^n \; v \quad \mapsto \mathrm{Span}\, v \tag{1.3}$$

is differentiable.

Let us show how to identify $\mathbb{R} \cup \{\infty\}$ to $\mathbb{R}P^1$, the set of directions in the plane $\mathbb{R}^2$. Indeed consider any vertical line l (which does contain the origin). A non vertical vector subspace of $\mathbb{R}^2$ intersects l in one and only one point, while the vertical vector subspace is associated with the infinite point.

1.1.2 Topology on manifold and submanifold

A subset $V \subset M$ is **open** if for every chart $\varphi\colon \mathcal{U} \to M$, the set $\varphi^{-1}(V \cap \varphi(\mathcal{U}))$ is open in $\mathcal{U}$.

Theorem 1.1.
This definition gives a topology on M which has the following properties:

(i) the charts maps are continuous,

(ii) the sets $\varphi_\alpha(\mathcal{U}_\alpha)$ are open.

Proof. First we prove that the open system defines a topology. For this, remark that φ_α^{-1} is injective (if not, there should be some multivalued points). Then $\varphi^{-1}(A \cap B) = \varphi^{-1}(A) \cap \varphi^{-1}(B)$. If V_1 and V_2 are open in M, then

$$\varphi^{-1}(V_1 \cap V_2 \cap \varphi(\mathcal{U})) = \varphi^{-1}(V_1 \cap \varphi(\mathcal{U})) \cap \varphi^{-1}(V_2 \cap \varphi(\mathcal{U}))$$

which is open in $\mathbb{R}^n$. The same property works for the unions.

Now we turn our attention to the continuity of $\varphi\colon \mathcal{U} \to M$; for an open set V in M, we have to show that $\varphi^{-1}(V)$ is open in $\mathcal{U} \subset \mathbb{R}^n$. But the definition of the topology on M, is precisely the fact that $\varphi^{-1}(V \cap \varphi(\mathcal{U}))$ is open. □

If M is a differentiable manifold and N, a subset of M, we say that N is a **submanifold** of dimension k if $\forall\, p \in N$, there exists a chart $\varphi\colon \mathcal{U} \to M$ around p such that

$$\varphi^{-1}(\varphi(\mathcal{U}) \cap N) = \mathbb{R}^k \cap \mathcal{U} := \{(x_1, \ldots, x_k, 0 \ldots, 0) \in \mathcal{U}\}.$$

In this case, N is itself a manifold of dimension k for which one can choose the φ of the definition as charts.

Let us consider M and N, two differentiable manifolds, $f\colon M \to N$ a C^∞ map and $x \in M$. We say that f is an **immersion** at x if $df_x\colon T_xM \to T_{f(x)}N$ is injective and that f is a **submersion** if df_x is surjective.

If M and M are two analytic manifolds, a map $\phi\colon M \to N$ is **regular** at $p \in M$ if it is analytic at p and $d\phi_p\colon T_pM \to T_{\phi(p)}N$ is injective.

Proposition 1.2.
Let M be a submanifold of the manifold N. If $p \in M$, then there exists a coordinate system $\{x_1, \ldots, x_n\}$ on a neighbourhood of p in N such that $x_1(p) = \ldots = x_n(p) = 0$ and such that the set

$$U = \{q \in V \;\; st \; x_j(q) = 0 \,\forall\, m+1 \leqslant j \leqslant n\}$$

gives a local chart of M containing p.

Proof. No proof. □

The sense of this proposition is that one can put p at the center of a coordinate system on N such that M is just a submanifold of N parametrised by the fact that its last $m-n$ components are zero.

Now we can give a characterization for a submanifold: N is a submanifold of M when $N \subset M$ (as set) and the identity $\iota\colon N \to M$ is regular.

Proposition 1.3.
The own topology of a submanifold is finer than the induced one from the manifold.

Proof. Let M be a manifold of dimension n and N a submanifold[2] of dimension $k < n$. We consider V, an open subset of N for the induced topology, so $V = N \cap \mathcal{O}$ for a certain open subset $\mathcal{O}$ of M. The aim is to show that V is an open subset in the topology of N.

Let us define $\mathcal{P} = \varphi^{-1}(\mathcal{O})$. The charts of N are the projection to $\mathbb{R}^k$ of the ones of M. We have to consider $W = \varphi^{-1}(V)$, since N is a submanifold, $\varphi^{-1}(\mathcal{O} \cap N) = \mathbb{R}^k \cap \mathcal{P}$. It is clear that $W = \mathbb{R}^k \cap \mathcal{P}$ is an open subset of $\mathbb{R}^k$ because it is the projection on the k first coordinates of an open subset of $\mathbb{R}^n$.

The subset V of N will be open in the sense of the own topology of N if $\varphi'^{-1}(V \cap \varphi'(\mathcal{U}'))$ is open in $\mathbb{R}^k$ where φ' is the restriction of φ to his k first coordinates: $\varphi'(a) = \varphi(a, 0)$ and $\mathcal{U}'$ is the projection of $\mathcal{U}$. □

Lemma 1.4.
Let V, M be two manifolds and $\varphi\colon V \to M$, a differentiable map. We suppose that $\varphi(V)$ is contained in a submanifold S of M. If $\varphi\colon V \to S$ is continuous, then it is differentiable.

Remark 1.5.
The map φ is certainly continuous as map from V to M (this is in the assumptions). But this don't imply that it is continuous for the topology on S (which is the induced one from M). So the continuity of $\varphi\colon V \to S$ is a true assumption.

Proof. Let $p \in V$. By proposition 1.2, we have a coordinate system $\{x_1, \ldots, x_m\}$ valid on a neighbourhood N of $\varphi(p)$ in M such that the set

$$\{r \in N \text{ st } x_j(r) = 0\, \forall s < j \leqslant m\}$$

with the restriction of $(x_1, \ldots x_s) \in N_S$ form a local chart which contains $\varphi(p)$. From the continuity of φ, there exists a chart (W, ψ) around p such that $\varphi(W) \subset N_S$. The coordinates $x_j(\varphi(q))$ are differentiable functions of the coordinates of q in W. In particular, the coordinates $x_j(\varphi(q))$ for $1 \leqslant j \leqslant s$ are differentiable and $\varphi\colon V \to S$ is differentiable because its expression in a chart is differentiable. □

A consequence of this lemma: if V and S are submanifolds of M with $V \subset S$, and if S has the induced topology from M, then V is a submanifold of S. Indeed, we can consider the inclusion $\iota\colon V \to S$: it is differentiable from V to M and continuous from V to S then it is differentiable from V to S by the lemma. Thus $V = \iota^{-1}(S)$ is a submanifold of S (this is a classical result of differential geometry).

Proposition 1.6.
A submanifold is open if and only if it has the same dimension as the main manifold.

Proof. Necessary condition. We consider some charts $\varphi_i\colon U_i \to M$ on some open subsets U_i of $\mathbb{R}^n$. If N is open in M, then this can be written as

$$N = \bigcup_i U_i.$$

If we choose the charts on M in such a manner that $\varphi_i\colon U_i \cap \mathbb{R}^k \to N$ are charts of N, we must have $\varphi_i(U_i \cap \mathbb{R}^k) = \varphi_i(U_i)$. Then it is clear that $k = n$ is necessary.
Sufficient condition. If N has same dimension as M, the charts $\varphi_i\colon U_i \to M$ are trivially restricted to N. □

1.2 Tangent and cotangent bundle

1.2.1 Tangent vector

As first attempt, we define a tangent vector of M at the point $x \in M$ as the "derivative" of a path $\gamma\colon (-\epsilon, \epsilon) \to M$ such that $\gamma(0) = x$. It is denoted by

$$\gamma'(0) = \frac{d}{dt}\gamma(t)\Big|_{t=0}.$$

[2] In the whole proof, we should say "there exists a sub-neighbourhood such that. . . "

The question is to correctly define de derivative in the right hand side. Such a definition is achieved as follows. A **tangent vector** to the manifold M is a linear map $X\colon C^\infty(M) \to \mathbb{R}$ which can be written under the form

$$Xf = (f \circ X)'(0) = \frac{d}{dt}\Big[f(X(t))\Big]_{t=0} \tag{1.4}$$

for a certain path $X\colon \mathbb{R} \to M$. Notice the abuse of notation between the tangent vector and the path which defines it.

A more formal way to define a tangent vector is to say that it is an equivalent class of path in the sense that two path are equivalent if and only if they induced maps by (1.4) are equals.

Using the chain rule $d(g \circ f)(a) = dg(f(a)) \circ df(a)$ for the differentiation in $\mathbb{R}^n$, one sees that this equivalence notion doesn't depend on the choice of φ. In other words, if φ and $\tilde\varphi$ are two charts for a neighbourhood of x, then $(\varphi^{-1} \circ \gamma)'(0) = (\varphi^{-1} \circ \sigma)'(0)$ if and only if $(\tilde\varphi^{-1} \circ \gamma)'(0) = (\tilde\varphi^{-1} \circ \sigma)'(0)$. The space of all tangent vectors at x is denoted by T_xM. There exists a bijection $[\gamma] \leftrightarrow (\varphi^{-1} \circ \gamma)'(0)$ between T_xM and $\mathbb{R}^n$, so T_xM is endowed with a vector space structure.

If $(\mathcal{U}, \varphi)$ is a chart around $X(0)$, we can express Xf using only well know objects by defining the function $\tilde f = f \circ \varphi$ and $\tilde X = \varphi^{-1} \circ X$

$$Xf = \frac{d}{dt}\Big[(\tilde f \circ \tilde X)(t)\Big]_{t=0} = \frac{\partial \tilde f}{\partial x^\alpha}\Big|_{x=\tilde X(0)} \frac{d\tilde X^\alpha}{dt}\Big|_{t=0}.$$

In this sense, we write

$$X = \frac{d\tilde X^\alpha}{dt}\frac{\partial}{\partial x^\alpha} \tag{1.5}$$

and we say that $\{\partial_1, \ldots, \partial_n\}$ is a basis of T_xM. As far as notations are concerned, from now a tangent vector is written as $X = X^\alpha \partial_\alpha$ where X^α is related to the path $X\colon \mathbb{R} \to M$ by $X^\alpha = d\tilde X^\alpha/dt$. We will no more mention the chart φ and write

$$Xf = \frac{d}{dt}\Big[f(X(t))\Big]_{t=0}.$$

Correctness of this short notation is because the equivalence relation is independent of the choice of chart. When we speak about a tangent vector to a given path $X(t)$ without specification, we think about $X'(0)$.

All this construction gives back the notion of tangent vector when $M \subset \mathbb{R}^m$. In order to see it, think to a surface in $\mathbb{R}^3$. A tangent vector is precisely given by a derivative of a path: if $c\colon \mathbb{R} \to \mathbb{R}^n$ is a path in the surface, a tangent vector to this curve is given by

$$\lim_{t\to 0}\frac{c(t_0) - c(t_0 + t)}{t}$$

which is a well know limit of a difference in $\mathbb{R}^n$.

Let us precise how does a tangent vector acts on maps others than $\mathbb{R}$-valued functions. If V is a vector space and $f\colon M \to V$, we define

$$Xf = (Xf^i)e_i$$

where $\{e_i\}$ is a basis of V and the functions $f^i\colon M \to \mathbb{R}$, the decomposition of f with respect to this basis. If we consider a map $\varphi\colon M \to N$ between two manifolds, the natural definition is $Xf := dfX$. More precisely, if we consider local coordinates x^α and a function $f\colon M \to \mathbb{R}$,

$$(d\varphi X)f = \frac{d}{dt}\Big[(f \circ \varphi \circ X)(t)\Big]_{t=0} = \frac{\partial f}{\partial x^\alpha}\frac{\partial \varphi^\alpha}{\partial x^\beta}\frac{dX^\beta}{dt}. \tag{1.6}$$

Now we are in a notational trouble: when we write $X = X^\alpha \partial_\alpha$, the "$X^\alpha$" is the derivative of the "X^α" which appears in the path $X(t) = (X^1(t), \ldots, X^n(t))$ which gives X by $X = X'(0)$. So equation (1.6) gives

$$X(\varphi) := d\varphi X = X^\beta(\partial_\beta \varphi^\alpha)\partial_\alpha. \tag{1.7}$$

1.2.2 Differential of a map

Let $f\colon M_1 \to M_2$ be a differentiable map, $x \in M_1$ and $X \in T_xM_1$, i.e. $X\colon \mathbb{R} \to M_1$ with $X(0) = x$ and $X'(0) = X$. We can consider the path $Y = f \circ X$ in M_2. The tangent vector to this path is written df_xX.

Proposition 1.7.
If $f\colon M_1 \to M_2$ is a differentiable map between two differentiable manifolds, the map

$$df_x\colon T_xM_1 \to T_{f(x)}M_2 \; X'(0) \quad\mapsto (f \circ X)'(0) \tag{1.8}$$

is linear.

Proof. We consider local coordinates $x\colon \mathbb{R}^n \to M_1$ and $y\colon \mathbb{R}^m \to M_2$. The maps $f\colon M_1 \to M_2$ and $y^{-1}\circ f\circ x\colon \mathbb{R}^n \to \mathbb{R}^m$ will sometimes be denoted by the same symbol f. We have $(x^{-1}\circ X)(t) = (x_1(t),\ldots,x_n(t))$ and $(y^{-1}\circ Y)(t) = \big(y_1(x_1(t),\ldots,x_n(t),\ldots,y_m(x_1(t),\ldots,x_n(t)\big)$, so that

$$Y'(0) = \left(\sum_{i=1}^n \frac{\partial y_1}{\partial x_i}x_i'(0),\ldots,\sum_{i=1}^n \frac{\partial y_m}{\partial x_i}x_i'(0)\right) \in \mathbb{R}^m$$

which can be written in a more matricial way under the form

$$Y'(0) = \left(\frac{\partial y_i}{\partial x_j}x_j'(0)\right).$$

So in the parametrisations x and y, the map df_x is given by the matrix $\partial y^i/\partial x_j$ which is well defined from the only given of f. □

Let $x\colon \mathcal{U} \to M$ and $y\colon \mathcal{V} \to M$ be two charts systems around $p \in M$. Consider the path $c(t) = x(0,\ldots,t,\ldots 0)$ where the t is at the position k. Then, with respect to these coordinates,

$$c'(0)f = \frac{d}{dt}\Big[f(c(t))\Big]_{t=0} = \frac{\partial f}{\partial x^i}\frac{dc^i}{dt} = \frac{\partial f}{\partial x^k},$$

so $c'(0) = \partial/\partial x^k$. Here, implicitly, we wrote $c^i = (x^i)^{-1}\circ c$ where $(x^i)^{-1}$ is the ith component of x^{-1} seen as element of $\mathbb{R}^n$. We can make the same computation with the system y. With these abuse of notation,

$$\frac{\partial}{\partial x^i} = \sum_j \frac{\partial y^j}{\partial x^i}\frac{\partial}{\partial y^j} \tag{1.9}$$

as it can be seen by applying it on any function $f\colon M \to \mathbb{R}$. More precisely if $x\colon \mathcal{U} \to M$ and $y\colon \mathcal{U} \to M$ are two charts (let $\mathcal{U}$ be the intersection of the domains of x and y), let $f\colon M \to \mathbb{R}$ and $\overline{f} = f\circ x$, $\tilde{f} = f\circ y$. The action of the vector ∂_{x^i} of the function f is given by

$$\partial_{x^i} f = \frac{\partial \overline{f}}{\partial x^i}$$

where the right hand side is a real number that can be computed with usual analysis on $\mathbb{R}^n$. This real *defines* the left hand side. Now, $\overline{f} = \tilde{f}\circ y^{-1}\circ x$, so that

$$\frac{\partial \overline{f}}{\partial x^i} = \frac{\partial(\tilde{f}\circ y^{-1}\circ x)}{\partial x^i} = \frac{\partial \tilde{f}}{\partial y^j}\frac{\partial y^j}{\partial x^i}$$

where $\frac{\partial \tilde{f}}{\partial y^j}$ is precisely what we write now by $\partial_{y^j} f$ and $\frac{\partial y^j}{\partial x^i}$ must be understood as the derivative with respect to x^i of the function $(y^{-1}\circ x)\colon \mathbb{R}^n \to \mathbb{R}^n$.

Let $f\colon M \to N$ and $g\colon N \to \mathbb{R}$; the definitions gives

$$(df_x X)g = \frac{d}{dt}\Big[(g\circ f)(X(t))\Big]_{t=0} = \frac{\partial g}{\partial y^i}\frac{\partial f^i}{\partial x^\alpha}\frac{dX^\alpha}{dt}.$$

This shows that $\frac{\partial f^i}{\partial x^\alpha}\frac{dX^\alpha}{dt}$ is $(df_x X)^i$. But dX^α/dt is what we should call X^α in the decomposition $X = X^\alpha\partial_\alpha$ then the matrix of df is given by $\frac{\partial f^i}{\partial x^\alpha}$. So we find back the old notion of differential.

Remark 1.8.
If $X \in T_xM$ and f is a vector valued *function on M, then one can define Xf by exactly the same expression. In this case,*

$$df X = \frac{d}{dt}\Big[f(v(t))\Big]_{t=0} = Xf.$$

A map $f\colon M_1 \to M_2$ is an **immersion** at $p \in M_1$ if $df_p\colon T_pM_1 \to T_{f(p)}M_2$ is injective. It is a **submersion** if df_p is surjective.

1.2.3 Tangent and cotangent bundle

If M is a n dimensional manifold, as set the tangent bundle is the *disjoint* union of tangent spaces

$$TM = \bigcup_{x\in M} T_xM.$$

Theorem 1.9.
The tangent bundle admits a $2n$ *dimensional manifold structure for which the projection*

$$\pi\colon TM \to M\ T_pM \quad \mapsto p \tag{1.10}$$

is a submersion.

The structure is easy to guess. If $\varphi_\alpha\colon \mathcal{U}_\alpha \to M$ is a coordinate system on M (with $\mathcal{U}_\alpha \subset \mathbb{R}^n$), we define $\psi_\alpha\colon \mathcal{U}_\alpha \times \mathbb{R}^n \to TM$ by

$$\psi(\underbrace{x_1,\ldots x_n}_{\in\mathcal{U}_\alpha}, \underbrace{a_1,\ldots a_n}_{\in\mathbb{R}^n}) = \sum_i a_i \left.\frac{\partial}{\partial x_i}\right|_{\varphi(x_1,\ldots,x_n)}.$$

The map $\psi_\beta^{-1}\circ\psi_\beta$ is differentiable because

$$(\psi_\beta^{-1}\circ\psi_\beta)(x,a) = (y(x), \sum_i a_i \left.\frac{\partial y_j}{\partial x_i}\right|_{y(x)})$$

which is a composition of differentiable maps. The set TM endowed with this structure is called the **tangent bundle**.

1.2.4 Vector space structure on the tangent space

If $X, Y \in T_pM$ are tangent vectors, one can define $X+Y$ and λX for every $\lambda\in\mathbb{R}$. The second one is easy:

$$\lambda X = \frac{d}{dt}\Big[X(\lambda t)\Big]_{t=0}. \tag{1.11}$$

In order to define the sum of two vectors one has to consider a neighbourhood $\mathcal{U}$ of p in M and a chart $\varphi\colon\mathcal{U}\to\mathcal{O}$ where $\mathcal{O}$ is an open set in $\mathbb{R}^n$. Then one consider a basis $\{e_i\}_{1\leqslant i\leqslant n}$ of $\mathbb{R}^n$ at the point $\varphi(p)$. With these choices we define the “basis” path

$$\gamma_i(t) = \varphi^{-1}(te_i) \tag{1.12}$$

and we write

$$\partial_i = \frac{\partial}{\partial x_i} = \frac{d}{dt}\Big[\varphi^{-1}(te_i)\Big]_{t=0}. \tag{1.13}$$

The vectors ∂_i form a basis of T_pM in the sense of the following lemma.

Lemma 1.10.
The action of a vector $X\in T_pM$ *on a function* $f\colon M\to\mathbb{R}$ *can be decomposed into*

$$Xf = \sum_{i=1}^n X_i(\partial_i f) \tag{1.14}$$

with $X_i\in\mathbb{R}$

Proof. Let $\varphi\colon M\to\mathbb{R}^n$ be a chart of a neighbourhood of p with $\varphi(p)=0$. We determine the value of X_i using the function

$$f_i(x) = \varphi(x)_i, \tag{1.15}$$

that is the ith component of the point $\varphi(x)\in\mathbb{R}^n$. Then if we write $\varphi\big(X(t)\big) = \sum_j a_j(t)e_j$ we have

$$X(f_i) = \frac{d}{dt}\Big[f_i\big(X(t)\big)\Big]_{t=0} = \frac{d}{dt}\Big[\big[\sum_j a_j(t)e_j\big]_i\Big]_{t=0} = \frac{d}{dt}\Big[ai_(t)\Big]_{t=0} = a_i'(0). \tag{1.16a}$$

Notice that $a_i(0)=0$ since $X(0)=p$ and $\varphi(p)=0$. The combination $f\circ\varphi^{-1}$ is an usual function from $\mathbb{R}^n$ to

$\mathbb{R}$, so that we can use the chain rule on it. The following computation thus make sense:

$$Xf = \frac{d}{dt}\Big[f\big(X(t)\big)\Big]_{t=0} \tag{1.17a}$$
$$= \frac{d}{dt}\Big[f\Big(\varphi^{-1}\varphi\big(X(t)\big)\Big)\Big]_{t=0} \tag{1.17b}$$
$$= \frac{d}{dt}\Big[(f\circ\varphi^{-1})\big(\sum_j a_j(t)e_j\big)\Big]_{t=0} \tag{1.17c}$$
$$= \sum_k \frac{\partial(f\circ\varphi^{-1})}{\partial x_k}\underbrace{\big(\sum_j a_j(0)e_j\big)}_{=\varphi(p)=0}\underbrace{\frac{d\big[\sum_j a_j(t)e_j\big]_k}{dt}}_{=a_k'(0)} \tag{1.17d}$$
$$= \sum_k a_k'(0)\frac{\partial(f\circ\varphi^{-1})}{\partial x_k}(0). \tag{1.17e}$$

Now using the definition of a derivative of a function $\mathbb{R}^n \to \mathbb{R}$ and of the "basis" tangent vector ∂_k,

$$\frac{\partial(f\circ\varphi^{-1})}{\partial x_k}(0) = \frac{d}{dt}\Big[(f\circ\varphi^{-1})(te_k)\Big]_{t=0} \tag{1.18a}$$
$$= \partial_k f \tag{1.18b}$$

At the end of the day we have

$$Xf = \sum_k a_k'(0)\partial_k f. \tag{1.19}$$

□

This lemma allows us to define the sum in T_pM as

$$\Big(\sum_k X_k\partial_k\Big) + \Big(\sum_k Y_k\partial_k\Big) = \sum_k (X_k+Y_k)\partial_k \tag{1.20}$$

when X_k and Y_k are reals.

The tangent space T_pM is thus a vector space.

1.2.5 Commutator of vector fields

If $X, Y \in \mathfrak{X}(M)$, one can define the **commutator** $[X,Y]$ in the following way. First remark that, if $f\colon M\to\mathbb{R}$, the object $X(f)$ is also a function from M to $\mathbb{R}$ by $X(f)(x) = X_x(f)$, so we can apply Y on $X(f)$. The definition of $[X,Y]_x$ is

$$[X,Y]_x f = X_x(Yf) - Y_x(Xf). \tag{1.21}$$

If $X = X^i\partial_i$ and $Y = Y^j\partial_j$, then $XY(f) = X^i\partial_i(Y^j\partial_j f) = X^i\partial_i Y^j\partial_j f + X^iY^j\partial^2_{ij}f$. From symmetry $\partial^2_{ij}f = \partial^2_{ji}f$, the difference $XYf - YXf$ is only $X^i\partial_iY^j - Y^i\partial_iX^j$, so that

$$[X,Y]^i = XY^i - YX^i \tag{1.22}$$

where X^i and Y^i are seen as functions from M to $\mathbb{R}$.

1.2.6 Some Leibnitz formulas

See [1], chapter I, proposition 1.4.

Lemma 1.11.
If M and N are two manifolds, we have a canonical isomorphism

$$T_{(p,q)}(M\times N) \simeq T_pM + T_qN.$$

Proof. A $Z\in T_{(p,q)}(M\times N)$ is the tangent vector to a curve $(x(t), y(y))$ in $M\times N$. We can consider $X\in T_pM$ given by $X = x'(0)$ and $Y\in T_qN$ given by $Y = y'(0)$. The isomorphism is the identification $(X,Y)\simeq Z$.

Indeed, let us define $\overline{X} \in T_{(p,q)}(M \times N)$, the tangent vector to the curve $(x(t), q)$, and $\overline{Y} \in T_{(p,q)}(M \times N)$, the tangent vector to the curve $(p, y(t))$. Then $Z = \overline{X} + \overline{Y}$ because for any $f\colon M \times N \to \mathbb{R}$,

$$Zf = \frac{d}{dt} f(x(t), y(t))\Big|_{t=0} = \frac{d}{dt} f(x(t), y(0))\Big|_{t=0} + \frac{d}{dt} f(x(0), y(t))\Big|_{t=0} = \overline{X}f + \overline{Y}f. \tag{1.23}$$

□

Proposition 1.12 (Leibnitz formula)**.**
Let us consider M, N, V, three manifold; a map $\varphi\colon M \times N \to V$ and a vector $Z \in T_{(p,q)}(M \times N)$ which corresponds (lemma 1.11) to $(X, Y) \in T_pM + T_qN$.

*If we define $\varphi_1\colon M \to V$ and $\varphi_2\colon N \to V$ by $\varphi_1(p') = \varphi(p', q)$ and $\varphi_2(q') = \varphi(p, q')$, we have the **Leibnitz formula**:*

$$d\varphi(Z) = d\varphi_1(X) + d\varphi_2(Y). \tag{1.24}$$

Proof. Since $Z = \overline{X} + \overline{Y}$, we just have to remark that

$$d\varphi(\overline{X}) = \frac{d}{dt}\varphi(x(t), q)\Big|_{t=0} = d\varphi_1(X),$$

so $d\varphi(Z) = d\varphi(\overline{X} + \overline{Y}) = d\varphi_1(X) + d\varphi_2(Y)$. □

One of the most important application of the Leibnitz rule is the corollary 4.6 on principal bundles.

1.2.7 Cotangent bundle

A form on a vector space V is a linear map $\alpha\colon V \to \mathbb{R}$. The set of all forms on V is denoted by V^* and is called the **dual space** of V. On each point of a manifold, one can consider the tangent bundle which is a vector space. Then one can consider, for each $x \in M$ the dual space $T_x^*M := (T_xM)^*$ which is called the **cotangent bundle**. A 1-**differential form** on M is a smooth map $\omega\colon M \to T^*M$ such that $\omega_x := \omega(x) \in T_x^*M$. So, for each $x \in M$, we have a 1-form $\omega_x\colon T_xM \to \mathbb{R}$.

Here, the smoothness is the fact that for any smooth vector field $X \in \mathfrak{X}(M)$, the map $x \to \omega_x(X_x)$ is smooth as function on M. One often considers vector-valued forms. This is exactly the same, but $\omega_x X_x$ belongs to a certain vector space instead of $\mathbb{R}$. The set of V-valued 1-forms on M is denoted by $\Omega(M, V)$ and simply $\Omega(M)$ if $V = \mathbb{R}$ The cotangent space T_p^*M of M at p is the dual space of T_pM, i.e. the vector space of all the (real valued) linear[3] 1-forms on T_pM. In the coordinate system $x\colon \mathcal{U} \to M$, we naturally use, on T_p^*M, the dual basis of the basis $\{\partial/\partial_{x^i}, \ldots \partial/\partial_{x^i}\}$ of T_pM. This dual basis is denoted by $\{dx_1, \ldots, dx_n\}$, the definition being as usual:

$$dx_i(\partial^j) = \delta_i^j. \tag{1.25}$$

The notation comes from the fact that equation (1.25) describes the action of the differential of the projection $x_i\colon \mathcal{U} \to \mathbb{R}$ on the vector ∂^j.

If $(\mathcal{U}_\alpha, \varphi_\alpha)$ is a chart of M, then the maps

$$\phi_\alpha\colon \mathcal{U}_\alpha \times \mathbb{R}^n \to T^*M \;\; (x, a) \quad \mapsto a^i dx_i|_x \tag{1.26}$$

give to T^*M a $2n$ dimensional manifold structure such that the canonical projection $\pi\colon T^*M \to M$ is an immersion.

When V is a finite-dimensional vector space, we denote by V^* its dual[4] and we often use the identifications $V \simeq V^* \simeq T_vV \simeq T_wV \simeq T_v^*V$ where v and w are any elements of V. Note however that there are no *canonical* isomorphism between these spaces, unless we consider some basis.

1.2.8 Exterior algebra

Here are some recall without proof about forms on vector space. If V is a vector space, we denote by $\Lambda^k V^*$ the space of all the k-form on V. We define $\wedge\colon \Lambda^k V^* \times \Lambda^l V^* \to \Lambda^{k+l} V^*$ by

$$(\omega^k \wedge \eta^l)(v_1, \ldots, v_{k+l}) = \frac{1}{k!l!} \sum_{\sigma \in S_{k+l}} sgn(\sigma)\omega(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\eta(v_{\sigma(k+1)}, v_{\sigma(k+1)}) \tag{1.27}$$

[3] When we say *a form*, we will always mean *a linear form*.
[4] The vector space of all the linear map $V \to \mathbb{R}$.

If $\{e_1, \ldots, e_n\}$ is a basis of V, the dual basis $\{\sigma^1, \ldots, \sigma^n\}$ of V^* is defined by $\sigma^i(e_j) = \delta^i_j$.

If $I = \{1 \leqslant i_1 \leqslant \ldots i_k \leqslant n\}$, we write $\sigma^I = \sigma^{i_1} \wedge \ldots \sigma^{i_k}$ any k-form can be decomposed as

$$\omega = \sum_I \omega_I \sigma^I.$$

The exterior algebra is provided with the **interior product** denoted by ι. It is defined by

$$\begin{aligned} \iota(v_0) \colon \Lambda^k W &\to \Lambda^{k-1} W \\ (\iota(v_0)\omega)(v_1, \ldots, v_{k-1}) &= \omega(v_0, v_1, \ldots, v_{k-1}). \end{aligned} \tag{1.28}$$

1.2.9 Pull-back and push-forward

Let $\varphi \colon M \to N$ be a smooth map, α a k-form on N, and Y a vector field on N. Consider the map $d\varphi \colon T_xM \to T_{\varphi(x)}M$. The aim is to extend it to a map from the tensor algebra of T_xM to the one of $T_{\varphi(x)}M$. See [1] for precise definition of the tensor algebra.

The **pull-back** of φ on a k-form α is the map

$$\varphi^* \colon \Omega^k(N) \to \Omega^k(M)$$

defined by

$$(\varphi^*\alpha)_m(v_1, \ldots, v_k) = \alpha_{\varphi(m)}(d\varphi_m v_1, \ldots, d\varphi_m v_k) \tag{1.29}$$

for all $m \in M$ and $v_i \in \mathfrak{X}(M)$.

Note the particular case $k = 0$. In this case, we take –instead of α– a function $f \colon N \to \mathbb{R}$ and the definition (1.29) gives $\varphi^* f \colon M \to \mathbb{R}$ by

$$\varphi^* f = f \circ \varphi.$$

The **push-forward** of φ on a k-form is the map

$$\varphi_* \colon \Omega^k(M) \to \Omega^k(N)$$

defined by $\varphi_* = (\varphi^{-1})^*$. For $v \in T_nN$, we explicitly have:

$$(\varphi_*\alpha)_n(v) = \alpha_{\varphi^{-1}(n)}(d\varphi_n^{-1} v).$$

Let now $\varphi \colon M \to N$ be a diffeomorphism. The **pull-back** of φ on a vector field is the map

$$\varphi^* \colon \mathfrak{X}(N) \to \mathfrak{X}(M)$$

defined by

$$(\varphi^* Y)(m) = [(d\varphi^{-1})_m \circ Y \circ \varphi](m),$$

or

$$(\varphi^* Y)_{\varphi^{-1}(n)} = (d\varphi^{-1})_n Y_n,$$

for all $n \in N$ and $m \in M$. Notice that

$$(d\varphi^{-1})_n \colon T_nN \to T_{\varphi^{-1}(n)}M,$$

and that $\varphi^{-1}(n)$ is well defined because φ is an homeomorphism.

The **push-forward** is, as before, defined by $\varphi_* = (\varphi^{-1})^*$. In order to show how to manipulate these notations, let us prove the following equation:

$$f_{*\xi} = (df)_\xi.$$

For $\varphi \colon M \to N$ and Y in $\mathfrak{X}(N)$, we just defined $\varphi^* \colon \mathfrak{X}(N) \to \mathfrak{X}(M)$, by

$$(\varphi^* Y)_{\varphi^{-1}(n)} = (d\varphi^{-1})_n Y_n. \tag{1.30}$$

Take $f \colon M \to N$; we want to compute $f_* = (f^{-1})^*$ with $(f^{-1})^* \colon \mathfrak{X}(M) \to \mathfrak{X}(N)$. Replacing the “$^{-1}$“ on the right places, the definition (1.30) gives us

$$\Big[(f^{-1})^* X\Big]_{f(m)} = (df)_m X_m,$$

if $X \in \mathfrak{X}(M)$, and $m \in M$.

We can rewrite it without any indices: the coherence of the spaces automatically impose the indices: $(f^{-1})^* X = (df)X$. It can also be rewritten as $(f^{-1})^* = df$, and thus $f_* = df$. From there to $f_{*\xi} = (df)_\xi$, it is straightforward.

1.2.10 Differential of k-forms

The differential of a k-form is defined by the following theorem.

Theorem 1.13.
Let M be a differentiable manifold. Then for each $k \in \mathbb{N}$, there exists an unique map

$$d\colon \Omega^k(M) \to \Omega^{k+1}(M)$$

such that

(i) d is linear,

(ii) for $k=0$, we find back the $d\colon C^\infty(M) \to \Omega^(M)$ previously defined,

(iii) if f is a function and ω^k a k-form, then

$$d(f\omega^k) = df \wedge \omega^k + f d\omega^k, \tag{1.31}$$

(iv) $d(\omega^k \wedge \eta^l) = d\omega^k \wedge \eta^l + (-1)^k \omega^k \wedge d\eta^l$,

(v) $d \circ d = 0$.

An explicit expression for $d\omega^k$ is actually given by

$$d\omega^k = \sum d\omega_I \wedge dx^I \tag{1.32}$$

if $\omega^k = \sum \omega_I dx^I$. An useful other way to write it is the following. If ω is a k-form and $X_1, \ldots, X_{p+1}$ some vector fields,

$$\begin{aligned}(k+1)d\omega(X_1, \ldots, X_{p+1}) &= \sum_{i=1}^{p+1} (-1)^{i+1} X_i \omega(X_1, \ldots \hat{X}_i, \ldots, X_{p+1}) \\ &\quad + \sum_{i<j} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \hat{X}_i, \hat{X}_j, \ldots, X_{p+1}).\end{aligned} \tag{1.33}$$

Let us show it with $p=1$. Let $\omega = \omega_i dx^i$ and compute $d\omega(X,Y) = \partial_i \omega_j (dx^i \wedge dx^j)(X,Y)$. For this, we have to keep in mind that the ∂_i acts only on ω_j while, in equation (1.33), a term $X\omega(Y)$ means –pointwise– the action of X on the function $\omega(Y)\colon M \to \mathbb{R}$. So we have to use Leibnitz formula:

$$(\partial_i \omega_j) X^i Y^j = (X\omega_j) Y^j = X(\omega_j Y^j) - \omega_j X Y^j.$$

On the other hand, we know that $[X,Y]^i = XY^i - YX^i$, so

$$d\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]). \tag{1.34}$$

1.2.10.1 Hodge operator

Let us take a manifold M endowed with a metric g. We can define a map $r\colon T_x^*M \to T_xM$ by, for $\alpha \in T_x^*M$,

$$\langle r(\alpha), v \rangle = \alpha(v).$$

for all $v \in T_xM$, where $\langle \cdot, \cdot \rangle$ stands for the product given by the metric g. If we have $\alpha, \beta \in T_x^*M$, we can define

$$\langle \alpha, \beta \rangle = \langle r(\alpha), r(\beta) \rangle.$$

With this, we define an inner product on $\Lambda^p(T_x^*M)$:

$$\langle \alpha_1 \wedge \ldots \alpha_p, \beta_1 \wedge \ldots \beta_p \rangle = \det_{ij} \langle \alpha_i, \beta_j \rangle.$$

The **Hodge operator** is $\star\colon \Lambda^p(T_x^*M) \to \Lambda^{n-p}(T_x^*M)$ such that for any $\phi \in \Lambda^p(T_x^*M)$,

$$\phi \wedge (\star\psi) = \langle \phi, \psi \rangle \Omega = \sqrt{|\det(g)|} dx^1 \wedge \ldots \wedge dx^n. \tag{1.35}$$

1.2.10.2 Volume form and orientation

Let M be a n dimensional smooth manifold. A **volume form** on M is a nowhere vanishing n-form and the manifold itself is said to be **orientable** if such a volume form exists. Two volume forms μ_1 and μ_2 are describe the same orientation if there exists a function $f > 0$ such that[5] $\mu_1 = f\mu_2$.

Proposition 1.14.
There exists only two orientations on a connected orientable manifold.

$\boxed{\textbf{Problem and misunderstanding}}$ **1.**
Check if the statement of that proposition is correct. Find a reference.

One says that the *ordered* basis $(v_1, \cdots, v_n)$ of T_xM is **positively oriented** with respect to the volume form μ is $\mu_x(v_1, \cdots, v_n) > 0$.

1.2.11 Musical isomorphism

In some literature, we find the symbols $v^\flat$ and $\alpha^\sharp$. What does it mean ? For $X \in \mathfrak{X}(M)$ and $\omega \in \Omega^2(M)$, the **flat** operation $v^\flat \in \Omega^1(M)$ is simply defined by the inner product:

$$v^\flat = i(v)\omega \tag{1.36}$$

In the same way, we define the **sharp** operation by taking a 1-form α and defining $\alpha^\sharp$ by

$$i(\alpha^\sharp)\omega = \alpha. \tag{1.37}$$

An immediate property is, for all $v \in \mathfrak{X}(M)$, $v^{\flat\sharp} = v$, and for all $\alpha \in \Omega^1(M)$, $\alpha^{\sharp\flat} = \omega$.

1.2.12 Lie derivative

Consider $X \in \mathfrak{X}(M)$ and $\alpha \in \Omega^p(M)$. Let $\varphi_t \colon M \to M$ be the flow of X. The **Lie derivative** of α is

$$\mathcal{L}_X\alpha = \lim_{t\to 0} \frac{1}{t}[(\varphi_t^*\alpha) - \alpha] = \frac{d}{dt}\varphi_t^*\alpha\Big|_{t=0}. \tag{1.38}$$

More explicitly, for $x \in M$ and $v \in T_xM$,

$$(\mathcal{L}_X\alpha)_x(v) = \lim_{t\to 0} \frac{1}{t}\left[(\varphi_t^*\alpha)_x(v) - \alpha_x(v)\right]$$

In the definition of the **Lie derivative** for a vector field, we need an extra minus sign:

$$(\mathcal{L}_XY)_x = \frac{d}{dt}\varphi_{-t*}Y_{\varphi_t(x)}\Big|_{t=0}. \tag{1.39}$$

Why a minus sign ? Because $Y_{\varphi_t(x)} \in T_{\varphi_t(x)}M$, but $(d\varphi_{-t})_a \colon T_aM \to T_{\varphi_{-t}(a)}M$ so that, if we want, $\varphi_{-t*}Y_{\varphi_t(x)}$ to be a vector at x, we can't use φ_{t*}.

These two definitions can be embedded in only one. Let $X \in \mathfrak{X}(M)$ and φ_t its integral curve[6]. We know that φ_{t*} is an isomorphism $\varphi_{t*} \colon T_{\varphi^{-1}(x)}M \to T_xM$. It can be extended to an isomorphism of the tensor algebras at $\varphi^{-1}(x)$ and x. We note it $\tilde{\varphi}_t$. For all tensor field K on M, we define

$$(\mathcal{L}_XK)_x = \lim_{t\to 0}[K_x - (\tilde{\varphi}_tK)_x].$$

On a Riemannian manifold (M, g), a vector field X is a **Killing vector field** if $\mathcal{L}_Xg = 0$.

Lemma 1.15.
Let $f \colon (-\epsilon, \epsilon) \times M \to \mathbb{R}$ be a differentiable map with $f(0, p) = 0$ for all $p \in U$. Then there exists $g \colon (-\epsilon, \epsilon) \times M \to \mathbb{R}$, a differentiable map such that $f(t, p) = tg(t, p)$ and

$$g(0, q) = \frac{\partial f(t, q)}{\partial t}\Big|_{t=0}.$$

[5] Recall that the space of n-forms is one-dimensional.
[6] *i.e.* for all $x \in M$, $\varphi_0(x) = x$ and $\frac{d}{dt}\varphi_{u+t}(x)\Big|_{t=0} = X_{\varphi_u(x)}$.

Proof. Take
$$g(t,q) = \int_0^1 \frac{\partial f(ts,p)}{\partial(ts)} ds,$$
and use the change of variable $s \to ts$. □

Lemma 1.16.
If φ_t is the integral curve of X, for all function $f\colon M \to \mathbb{R}$, there exists a map g, $g_t(p) = g(t,p)$ such that $f \circ \varphi_t = f + tg_t$ and $g_0 = Xf$.

Proof. Consider $\overline{f}(t,p) = f(\varphi_t(p)) - f(p)$, and apply the lemma:
$$f \circ \varphi_t = tg_t(p) + f(p).$$
Thus we have
$$Xf = \lim_{t\to 0} \frac{1}{t}[f(\varphi_t(p)) - f(p)] = \lim_{t\to 0} g_t(p) = g_0(p).$$
□

One of the main properties of the Lie derivative is the following:

Theorem 1.17.
Let X, $Y \in \mathfrak{X}(M)$ and φ_t be the integral curve of X. Then
$$[X,Y]_p = \lim_{t\to 0} \frac{1}{t}[Y - d\varphi_t Y](\varphi_t(p)),$$
or
$$\mathcal{L}_X Y = [X,Y].$$

Proof. Take $f\colon M \to \mathbb{R}$ and the function given by the lemma: $g_t\colon M \to \mathbb{R}$ such that $f \circ \varphi_t = f + tg_t$ and $g_0 = Xf$. Then put $p(t) = \varphi_t^{-1}(p)$. The rest of the proof is a computation:
$$(\varphi_{t*}Y)_p f = Y(f \circ \varphi_t)_{p(t)} = (Yf)_{p(t)} + t(Yg_t)_{p(t)},$$
so
$$\begin{aligned}
\lim_{t\to 0} \frac{1}{t}[Y_p - (\varphi_{t*}Y)_p]f &= \lim_{t\to 0} \frac{1}{t}[(Yf)_p - (Yf)_{p(t)}] - \lim_{t\to 0}(Yg_t)_{p(t)} \\
&= X_p(Yf) - Y_p g_0 \\
&= [X,Y]_p f.
\end{aligned} \tag{1.40}$$
□

A second important property is

Theorem 1.18.
For any function $f\colon M \to V$,
$$\mathcal{L}_X f = Xf.$$

Proof. If $X(t)$ is the path which defines the vector X, it is obvious that at $t = 0$, $X(t)$ is an integral curve to X, so that we can take $X(t)$ instead of φ_t in (1.38). Therefore we have:
$$\mathcal{L}_X f = \frac{d}{dt}\varphi_t^* f\Big|_{t=0} = Xf \tag{1.41}$$
by definition of the action of a vector on a function.
□

Chapter 2

Lie groups and subgroups

Most of this chapter come (often very directly) from [3]. Other sources are [6–9].

Do you know what is violet and commutative ? Answer in the footnote[1].

2.1 Lie groups

A **Lie group** is a manifold G endowed with a group structure such that the inversion map $i\colon G \to G$, $i(x) = x^{-1}$ and the multiplication $m\colon G \times G \to G$, $m(x,y) = xy$ are differentiable. The **Lie algebra** of the Lie group G is the tangent space of G at the identity: $\mathcal{G} = T_eG$.

It is immediate to see that $g \mapsto g^{-1}$ is a smooth homeomorphism and that, for any fixed g_0, g_1, the maps

$$\begin{aligned} g &\mapsto g_0 g, \\ g &\mapsto g g_0, \\ g &\mapsto g_0 g g_1 \end{aligned}$$

are smooth homeomorphisms. When $A \subset G$, we define $A^{-1} = \{g^{-1} \text{ st } g \in G\}$.

2.1.1 Connected component of Lie groups

Proposition 2.1.
If G is a connected Lie group and $\mathcal{U}$, a neighbourhood of the identity e, then G is generated by $\mathcal{U}$ in the sense that $\forall g \in G$, there exists a finite *number of $g_i \in \mathcal{U}$ such that*

$$g = g_1 \ldots g_n.$$

Notice that the number n is function of g in general.

Proof. Eventually passing to a subset, we can suppose that $\mathcal{U}$ is open. In this case, $\mathcal{U}^{-1}$ is open because it is the image of $\mathcal{U}$ under the homeomorphism $g \mapsto g^{-1}$. Now we consider $V = \mathcal{U} \cap \mathcal{U}^{-1}$. The main property of this set is that $V = V^{-1}$. Let

$$[V] = \{g_1 \ldots g_n \text{ st } g_i \in V\};$$

we will prove that $[V] = G$ by proving that it is closed and open in G (the fact that G is connected then concludes).

We begin by openness of $[V]$. Let $g_0 = g_1 \cdots g_n \in [V]$. We know that g_0V is open because the multiplication by g_0 is an homeomorphism. It is clear that $g_0V \subset [V]$ and that $g_0 = g_0e \in g_0V$. Hence $g_0 \in g_0V \subseteq [V]$. It proves that $[V]$ is open because g_0V is a neighbourhood of g_0 in $[V]$.

We now turn our attention to the closeness of $[V]$. Let $h \in \overline{[V]}$. The set hV is an open set which contains h and $hV \cap [V] \neq \varnothing$ because an open which contains an element of the closure of a set intersects the set (it is almost the definition of the closure). Let $g_0 \in hV \cap [V]$. There exists a $h_1 \in V$ such that $g_0 = hh_1$. For this h_1, we have $hh_1 = g_0 = g_1 \cdots g_n$, and therefore

$$h = g_1 \cdots g_n h_1^{-1} \in [V].$$

This proves that $h \in [V]$ because $h_1^{-1} \in V$ from the fact that $V = V^{-1}$. □

Remark that this proof emphasises the topological aspect of a Lie group: the differential structure was only used to prove thinks like that A^{-1} is open when A is open.

[1] An abelian grape !

Proposition 2.2.
Let G be a Lie group and G_0, the identity component of G. We have the following:

(i) G_0 is an open invariant subgroup of G,

(ii) G_0 is a Lie group,

(iii) the connected components of G are lateral classes of G_0. More specifically, if x belongs to the connected component G_1, then $G_1 = xG_0 = G_0x$.

Proof. We know that when M_1 is open in the manifold M, one can put on M_1 a differential structure of manifold of same dimension as M with the induced topology. Since G_0 is open, it is a smooth manifold. In order for G_0 to be a Lie group, we have to prove that it is stable under the inversion and that $gh \in G_0$ whenever $g, h \in G_0$.

First, G_0^{-1} is connected because it is homeomorphic to G_0 in G. The element e belongs to the intersection of G_0 and G_0^{-1}, so $G_0 \cup G_0^{-1}$ is connected as non-disjoint union of connected sets. Hence $G_0 \cup G_0^{-1} = G_0$ and we conclude that $G_0^{-1} \subseteq G_0$. The set G_0G_0 is connected because it is the image of $G_0 \times G_0$ under the multiplication map, but $e \in G_0G_0$, so $G_0G_0 \subseteq G_0$ and G_0 is thus closed for the multiplication. Hence G_0 is a Lie group.

For all $x \in G$, we have $e = xex^{-1} \in xG_0x^{-1}$, but xG_0x^{-1} is connected. Hence $xG_0x^{-1} \subseteq G_0$, which proves that G_0 is an invariant subset of G.

Lateral classes xG_0 are connected because the left multiplication is an homeomorphism. They are moreover *maximal* connected subsets because, if $xG_0 \subset H$ (proper inclusion) with a connected H, then $G_0 \subset x^{-1}H$ (still proper inclusion). But the definition of G_0 is that this proper inclusion is impossible. Therefore, the sets of the form xG_0 are maximally connected sets. It is clear that $\cup_{g\in G} gG_0 = G$.

Notice that the last point works with G_0x too. □

2.2 Two words about Lie algebra

2.2.1 The Lie algebra of $\mathrm{SU}(n)$

Let consider $G = \mathrm{SU}(n)$; the elements are complexes $n \times n$ matrices U such that $UU^\dagger = \mathbb{1}$ and $\det U = 1$. An element of the Lie algebra is given by a path $u\colon \mathbb{R} \to G$ in the group with $u(0) = \mathbb{1}$. Since for all t, $u(t)u(t)^\dagger = \mathbb{1}$,

$$\begin{aligned} 0 &= \frac{d}{dt}\Big[u(t)u(t)^\dagger\Big]_{t=0} \\ &= u(0)\frac{d}{dt}\Big[u(t)^\dagger\Big]_{t=0} + \frac{d}{dt}\Big[u(t)\Big]_{t=0} u(0)^\dagger \\ &= [d_t u(t)]^\dagger + [d_t u(t)]. \end{aligned} \tag{2.1}$$

So a general element of the Lie algebra $\mathfrak{su}(n)$ is an anti-hermitian matrix.

An element of $\mathrm{SU}(n)$ has also a determinant equal to 1. What condition does it implies on the elements of the Lie algebra ? If $g(t)$ is a path in $\mathrm{SU}(n)$ with $g(0) = \mathbb{1}$ we have

$$\det \begin{pmatrix} g_{11}(t) & g_{12}(t) & \ldots \\ f_{21}(t) & g_{22}(t) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = g_{11}(t)M_{11}(t) + g_{12}(t)M_{12}(t) + \ldots = 1 \tag{2.2}$$

where M_{ij} is the minor of g. If we derive the left hand side we get

$$g'_{11}(0)M_{11}(0) + g_{11}(0)M'_{11}(0) + g'_{12}(0)M_{12}(0) + g_{12}(0)M'_{12}(0) + \ldots \tag{2.3}$$

where the numbers $g'_{ij}(0)$ are the matrix entries of the tangent matrix, that is the matrix elements of a general element in $\mathfrak{su}(n)$. Since $g(0) = \mathbb{1}$ we have $M_{11}(0) = 1$, $g_{11}(0) = 1$, $M_{12}(0) = 0$ and $g_{12}(0) = 0$. Thus we have

$$(\det g)'(0) = X_{11} + M'_{11}(0) \tag{2.4}$$

where $X = g'(0)$. By induction we found that the trace of X appears. Thus the elements of $\mathfrak{su}(n)$ have vanishing trace.

2.2.2 What is $g^{-1}dg$?

The expression $g^{-1}dg$ is often written in the physical literature. In our framework, the way to gives a sense to this expression is to consider it pointwise acting on a tangent vector. More precisely, the framework is the data

of a manifold M, a Lie group G and a map $g\colon M \to G$. Pointwise, we have to apply $g(x)^{-1}dg_x$ to a tangent vector $v \in T_xM$.

Note that $dg_x : T_xM \to T_{g(0)}G \neq T_eG$, so $dg_x \notin \mathcal{G}$. But the product $g(x)^{-1}dg_xv$ is defined by

$$g(x)^{-1}dg_xv = \frac{d}{dt}\Big[g(x)^{-1}g(v(t))\Big]_{t=0} \in \mathcal{G}. \tag{2.5}$$

2.2.3 Invariant vector fields

If G is a Lie group, a vector field $\tilde{X} \in \Gamma^{\infty}(TG)$ is **left invariant** if

$$(dL_g)\tilde{X} = \tilde{X}. \tag{2.6}$$

To each element X in the Lie algebra $\mathcal{G}$, we have an associated left invariant vector field given on the point $g \in G$ by the path

$$\tilde{X}_g(t) = ge^{tX}. \tag{2.7}$$

In the same way, we associate a **right invariant** given by

$$\utilde{X}_g(t) = e^{tX}g \tag{2.8}$$

The invariance means that $(dL_h)_g\tilde{X}_g = \tilde{X}_{hg}$ and $(dR_h)_g\utilde{X}_g = \utilde{X}_{gh}$. The invariant vector fields are important because they carry the structure of the tangent space at identity (the Lie algebra). More precisely we have the following result:

Theorem 2.3.
The map $X \to X_e$ is a bijection between the left invariant vector fields on a Lie group and its Lie algebra T_eG.

Invariant vector fields are also often used in order to transport a structure from the identity of a Lie group to the whole group by $A_g(X_g) = A_e(dL_{g^{-1}}X_g)$ where A_e is some structure and X_g, a vector at g.

Proposition 2.4.
Let G be a Lie group and $\mathcal{G}$ the vector space of its left invariant vector fields.

(i) The map

$$\begin{aligned} \mathcal{G}\colon &\to T_eG \\ \tilde{X} &\mapsto \tilde{X}_e \end{aligned} \tag{2.9}$$

is a vector space isomorphism.

(ii) We have $[\mathcal{G},\mathcal{G}] \subset \mathcal{G}$ and $\mathcal{G}$ is a Lie algebra. Here, the commutator is the bracket of vector fields.

Proof. No proof. □

2.2.4 Invariant vector field

A vector field X on a Lie group G is **left invariant** if $dL_g(X) = X$ for every $g \in G$. Here $L_g\colon G \to G$ is the left translation defined by $L_g(h) = gh$. More explicitly, the left invariance is expressed by

$$\frac{d}{dt}\Big[gX_h(t)\Big]_{t=0} = X_{gh} \tag{2.10}$$

where $X_h(t)$ is the path defining the tangent vector $X_h \in T_hG$.

We want to prove that the vector space of left invariant vector fields is isomorphic to the tangent vector space T_eG to G at identity. If $X \in T_eG$, we introduce the left invariant vector field $X^L = dLX$, more explicitly:

$$X^L_g = \frac{d}{dt}\Big[gX(t)\Big]_{t=0}. \tag{2.11}$$

Then we consider $\alpha_X : I \to G$ the integral curve of maximal length to X^L trough X_e. Here, I is the interval on which α_X is defined. This is the solution of

$$\begin{cases} \dfrac{d}{dt}\Big[\alpha_X(t_0+t)\Big]_{t=0} = X_{\alpha_X(t_0)} & (2.12a) \\ \alpha_X(0) = e. & (2.12b) \end{cases}$$

Proposition 2.5.
Let $X \in T_eG$. The integral curve has $\mathbb{R}$ as domain and for every $s, t \in \mathbb{R}$,

$$\alpha_X(s+t) = \alpha_X(s)\alpha_X(t). \tag{2.13}$$

Proof. Let α be any integral curve for X^L and $y \in G$. If we put $\alpha_1(t) = y\alpha(t)$, we have

$$\frac{d}{dt}\Big[\alpha_1(t)\Big]_{t=0} = X^L_y, \tag{2.14}$$

so that α_1 is an integral curve for X^L trough the point y.

Let now I be the maximal domain of α_X, and $t_1 \in I$. If we set $x_1 = \alpha_X(t_1)$, the path

$$\alpha_1(t) = x_1\alpha_X(t) \tag{2.15}$$

is an integral curve of X^L trough x_1 and has the same maximal definition domain I. On the other hand, the maximal integral curve starting at e being α_X, the maximal integral curve starting at $\alpha_X(t_1)$ is

$$\alpha_2 \colon t \mapsto \alpha_X(t+t_1). \tag{2.16}$$

Its domain is $I - t_1$, but since it starts at x_1, it has to be the same as α_1, then $I \subset I - t_1$ which proves that $I = \mathbb{R}$.

For each s and t in $\mathbb{R}$, the maximal integral curve starting at $\alpha_X(s)$ can be written as

$$c(t) = \alpha_X(s)\alpha_X(t) \tag{2.17}$$

as well as

$$d(t) = \alpha_X(s+t), \tag{2.18}$$

so again by unicity, $\alpha_X(s+t) = \alpha_X(s)\alpha_X(t)$. □

2.2.5 Integral curve and exponential

If α_X is the integral curve to X^L, we define the **exponential**

$$\begin{aligned} \exp \colon T_eG &\to G \\ X &\mapsto \alpha_X(1). \end{aligned} \tag{2.19}$$

This definition works on Lie groups thanks to the group structure that allows to build a natural vector field X^L from the data of a single vector X. On general manifolds, one has not a notion of exponential. However, if one has a Riemannian manifold, one consider the geodesic.

In the case of groups for which the Killing form defines a scalar product, the notion of exponential associated with the Riemannian structure propagated from the Killing form coincides with the definition (2.19).

2.2.6 Adjoint map

The ideas of this short note comes from [10]. A more traumatic definition of the adjoint group can be found in [3], chapter II, §5. Let G be a Lie group, and $\mathcal{G}$, its Lie algebra. We define the **adjoint map** at the point $x \in G$ by

$$\begin{aligned} \mathbf{Ad}_x \colon G &\to G \\ \mathbf{Ad}_x\, y &= xyx^{-1} \end{aligned} \tag{2.20}$$

Then we define

$$Ad_x := (d\,\mathbf{Ad}_x)_e \colon \mathcal{G} \to \mathcal{G};$$

the chain rule applied on $\mathbf{Ad}_{xy} = \mathbf{Ad}_x \circ \mathbf{Ad}_y$ leads to $Ad_{xy} = Ad_x \circ Ad_y$, and thus we can see Ad as a group homomorphism $Ad \colon G \to GL(\mathcal{G})$, $Ad(x) = Ad_x$.

Definition 2.6.
*This homomorphism is the **adjoint representation** of the group G in the vector space $\mathcal{G}$.*

Finally, we define

$$ad := d(Ad)_1 \colon \mathcal{G} \to L(\mathcal{G}, \mathcal{G})$$

where we identify $T_1GL(\mathcal{G})$ with $L(\mathcal{G}, \mathcal{G})$.

Lemma 2.7.
If $f\colon G\to G$ is an automorphism of G (i.e.: $f(xy)=f(x)f(y)$), then df_e is an automorphism of $\mathcal{G}$: $df[X,Y]=[df X, df Y]$

Proof. First, remark that $f(\mathbf{Ad}_x\, y)=\mathbf{Ad}_{f(x)}\, f(y)$. Now, $\mathrm{Ad}_x\, X=(d\,\mathbf{Ad}_x)_e X$, so that one can compute:

$$\begin{aligned} df(\mathrm{Ad}_x\, X) &= \frac{d}{dt}\Big[f(\mathbf{Ad}_x\, X(t))\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\,\mathbf{Ad}_{f(x)}\, f(X(t))\Big]_{t=0}\\ &= (d\,\mathbf{Ad}_{f(x)})_{f(e)} df X\\ &= \mathrm{Ad}_{f(x)}\, df X. \end{aligned}\tag{2.21}$$

On the other hand, we need to understand how does the ad work.

$$\mathrm{ad}\, XY=\frac{d}{dt}\Big[\,\mathrm{Ad}_{X(t)}\,\Big]_{t=0}Y=\frac{d}{dt}\Big[\,\mathrm{Ad}_{X(t)}\, Y\Big]_{t=0}$$

because $\mathrm{Ad}_{X(t)}\colon\mathcal{G}\to\mathcal{G}$ is linear, so that Y can enter the derivation (for this, we identify $\mathcal{G}$ and $T_X\mathcal{G}$). Since $\mathrm{Ad}_{X(t)}\, Y$ is a path in $\mathcal{G}$ the *true space* is

$$(\mathrm{ad}\, X)Y=\frac{d}{dt}\Big[\,\mathrm{Ad}_{X(t)}\, Y\Big]_{t=0}\in T_{[X,Y]}\mathcal{G}\simeq\mathcal{G}.$$

For the same reason of linearity, df can get in the derivative in the expression $df\frac{d}{dt}\Big[\,\mathrm{Ad}_{X(t)}\, Y\Big]_{t=0}$. Thus

$$\begin{aligned} (\mathrm{ad}\, X)Y &= \frac{d}{dt}\Big[df\big(\,\mathrm{Ad}_{X(t)}\, Y\big)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\,\mathrm{Ad}_{f(X(t))}\, dfY\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\,\mathrm{Ad}_{f(X(t))}\,\Big]_{t=0}dfY\\ &= \mathrm{ad}(dfX)dfY\\ &= [dfX, dfY] \end{aligned}\tag{2.22}$$

because $f(X(t))$ is a path which gives dfX.

$\square$

One can show that $[X,Y]$ is tangent to the curve

$$c(t)=e^{-\sqrt{s}X}e^{-\sqrt{s}Y}e^{\sqrt{s}X}e^{\sqrt{s}Y}.\tag{2.23}$$

Lemma 2.8.
In the case of Lie algebra, the bracket is given by the derivative of the adjoint action:

$$\frac{d}{dt}\Big[\,\mathrm{Ad}(e^{tX})Y\Big]_{t=0}=[X,Y]\tag{2.24}$$

Proof. Let us make $[\tilde X,\tilde Y]_e$ act on a function f. Using the definition (1.39) and the property of theorem 1.17, we have

$$\begin{aligned} [\tilde X,\tilde Y]_e f &= \frac{d}{dt}\Big[(d\varphi^X_{-t})\tilde Y\Big]_{t=0}f\\ &= \frac{d}{dt}\Big[(d\varphi^X_{-t})_{\varphi^X_t(e)}\big(\tilde Y_{\varphi^X_t(e)}\big)\Big]_{t=0}f\\ &= \frac{d}{dt}\Big[\tilde Y_{e^{tX}}\cdot(f\circ\varphi^X_{-t})\Big]_{t=0} \end{aligned}\tag{2.25}$$

Now, we use the fact that, by definition, $\varphi^X_t(x)=xe^{tX}$, so that $\varphi^Y_s(e^{tX})=e^{tX}e^{sY}$ and we get

$$\begin{aligned} [\tilde X,\tilde Y]_e f &= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f\big(\varphi^X_{-t}(e^{tX}e^{sY})\big)\Big]_{s=0}\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f(e^{tX}e^{sY}e^{-tX})\Big]_{s=0}\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\frac{d}{ds}\Big[f\big(e^{s\,\mathrm{Ad}(e^{tX})Y}\big)\Big]_{s=0}\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\big(\,\mathrm{Ad}(e^{tX})Y\big)_e\cdot f\Big]_{t=0} \end{aligned}\tag{2.26}$$

$\square$

2.3 Fundamental vector field

If $\mathcal{G}$ is the Lie algebra of a Lie group G acting on a manifold M (the action of g on x being denoted by $x \cdot g$), the **fundamental vector field** associated with $A \in \mathcal{G}$ is given by

$$A^*_x = \frac{d}{dt}\Big[x \cdot e^{-tA}\Big]_{t=0}. \tag{2.27}$$

We always suppose that the action is effective. If the action of G is transitive, the fundamental vectors at point $x \in M$ form a basis of T_xM. More precisely, we have the

Lemma 2.9.
For any $v \in T_xM$, there exists a $A \in \mathcal{G}$ such that $v = A^*_x$, in other terms

$$\mathrm{Span}\{A^*_x \;\, st\;\, A \in \mathcal{G}\} = T_xM.$$

Proof. The vector v is given by a path $v(t)$ in M. Since the action is transitive, one can write $v(t) = x \cdot c(t)$ for a certain path c in G which fulfills $c(0) = e$. We have to show that v depends only on $c'(0) \in \mathcal{G}$. We consider

$$\begin{aligned} R\colon G \times M &\to M \\ R(g,x) &= x \cdot g, \end{aligned} \tag{2.28}$$

so

$$v = \frac{d}{dt}\Big[R(c(t),x)\Big]_{t=0} = dR_{(e,x)}\big[(d_t c(t),x) + (c(0),x)\big]. \tag{2.29}$$

□

Lemma 2.10.
If A, $B \in \mathcal{G}$ are such that $A^* = B^*$, and if the action is effective, then $A = B$.

Proof. We consider once again the map (2.28) and we look at

$$v = \frac{d}{dt}\Big[R(c(t),x)\Big]_{t=0} = (dR)_{(e,x)}\frac{d}{dt}\Big[(c(t),x)\Big]_{t=0},$$

keeping in mind that $c(t) = e^{-tA}$. In order to treat this expression, we define

$$R_1\colon G \to M, \quad R_1(h) = R(h,x), \tag{2.30a}$$

$$R_2\colon M \to M, \quad R_2(y) = R(g,y). \tag{2.30b}$$

So

$$v = dR_1(X) + dR_2(0) = dR_1c'(0)$$

and the assumption $A^*_x = B^*_x$ becomes $dR_1A = dR_1B$. This makes, for small enough t, $R_1(e^{tA}e^{-tB}) = x \cdot e^{tA}e^{-tB} = x$; if the action is effective, it imposes $A = B$. □

Lemma 2.11.
If we consider the action of a matrix group, R_g acts on the fundamental field by

$$dR_g(A^*_\xi) = \big(\mathrm{Ad}(g^{-1})A\big)^*_{\xi\cdot g}.$$

Proof. Just notice that $e^{-t\,\mathrm{Ad}(g^{-1})A} = \mathbf{Ad}_{g^{-1}}(e^{-tA}) = g^{-1}e^{-tA}g$, thus

$$\big(\mathrm{Ad}(g^{-1})A\big)^*_{\xi\cdot g} = \frac{d}{dt}\Big[\xi \cdot ge^{-t\,\mathrm{Ad}(g^{-1})A}\Big]_{t=0} = dR_g(A^*_\xi). \tag{2.31}$$

□

2.4 Exponential map

A **topological group** is a group G equipped with a topological structure such that the maps $(x, y) \in G^2 \to xy \in G$ and $x \in G \to x^{-1} \in G$ are continuous.

Remark 2.12.
From the existence of an unique inverse for any element of G, the multiplication and the inversion are also open maps.

A **Lie group** is a group G which is in the same times an analytic manifold such that the group operations (multiplication and inverse) are analytic. In particular, we *do not* suppose that they are diffeomorphism.

The concept of normal neighbourhood will be widely used for the study of the relations between a Lie group and its algebra. Let M be a differentiable manifold. If V is a neighbourhood of zero in T_pM on which the exponential $\exp_p \colon T_pM \to M$ is a diffeomorphism, then $\exp_p V$ is **normal neighbourhood** of p.

Lemma 2.13.
Let G, H be two Lie groups with algebras $\mathcal{G}$ and $\mathcal{H}$. Let $\phi \colon G \to H$ be a homomorphism differentiable at e, the unit in G. Then for all $X \in \mathcal{G}$, the following formula holds:

$$\phi(\exp X) = \exp(d\phi_e X).$$

It can be found in [10].

Proposition 2.14.
Let G be a connected Lie group.

(i) All the left invariant vector fields are complete. That means that the map $X \mapsto e^X$ is defined for every $X \in \mathcal{G}$.

(ii) The map $\exp \colon \mathcal{G} \to G$ is a local diffeomorphism in a neighbourhood of 0 in $\mathcal{G}$.

Proof. (i) The flow is a one parameter subgroup. Thus if e^{tX} is defined for $t \in [0, a]$, by composition, e^{2a} is defined. So e^{tX} is defined for every value of t in $\mathbb{R}$.

(ii) Let us consider the manifold $G \times \mathcal{G}$ and the vector field Ξ defined by

$$\Xi_{(g,X)} = \tilde{X}_g \oplus 0 \in T_g(G) \oplus T_X\mathcal{G} \simeq T_{(g,X)}(G \times \mathcal{G}). \tag{2.32}$$

The flow of that vector field is given by

$$\Phi_t(g, X) = \big(g \exp(tX), X\big). \tag{2.33}$$

In particular, Ξ is a complete vector field, and we consider the global diffeomorphism

$$\begin{aligned} \Phi_1 \colon G \times \mathcal{G} &\to G \times \mathcal{G} \\ (g, X) &\mapsto \big(g \exp(X), X\big). \end{aligned} \tag{2.34}$$

On the point (e, X) we have $\Phi_1(e, X) = (\exp(X), X)$. Thus the exponential is the projection on the first component of $\Phi_1(e, X)$ and we can write

$$\exp(X) = \mathtt{pr}_1 \circ \Phi_1(e, X). \tag{2.35}$$

It is a smooth function since both the projection and Φ_1 are smooth.

Now, the differential $(d\exp)_0$ is the identity on $\mathcal{G}$, so that the theorem of inverse function makes exp a local diffeomorphism.

□

Theorem 2.15.
For any $p \in M$, there exist a $\delta > 0$ and a neighbourhood W of p in M such that for every $q \in W$, we have

- $\exp_q$ *is a diffeomorphism on $B_\delta(0) \subset T_qM$,*
- $\exp_q B_\delta(0)$ *contains W*

This theorem says that everywhere on a differentiable manifold, one can find a neighbourhood which is a normal neighbourhood of each of its points. Such a neighbourhood is said a *totally* normal neighbourhood.

Lemma 2.16.
In a Lie group, e is an isolated fixed point for the inversion.

Proof. One can use an exponential map in a neighbourhood of e. In this neighbourhood, an element g can be written as $g = e^X$ for a certain $X \in \mathfrak{g}$. The equality $g = g^{-1}$ gives (because the exponential is a diffeomorphism) $X = -X$, so that $X = 0$ and $g = e$. □

Lemma 2.17.
Let $\mathfrak{g}$ be a Lie algebra and A, a linear operator on $\mathfrak{g}$ (see as a common vector space) such that $\forall t \in \mathbb{R}$, the map e^{tA} is an automorphism of $\mathfrak{g}$. Then A is a derivation of $\mathfrak{g}$.

Proof. Let us consider $X, Y \in \mathfrak{g}$; the assumption is

$$e^{tA}[X,Y] = [e^{tA}X, e^{tA}Y].$$

Since e^{tA} is a linear map, it has a "good behavior" with the derivations:

$$\frac{d}{dt}\Big(e^{tA}[X,Y]\Big)_{t=0} = \frac{d}{dt}\Big(e^{tA}\Big)_{t=0}[X,Y] = A[X,Y].$$

Using on the other hand the linearity of ad, we can see

$$(\mathrm{ad}(e^{tA}X))(e^{tA}Y)$$

as a product "matrix times vector". Then

$$\begin{aligned}\frac{d}{dt}\Big([e^{tA}X, e^{tA}Y]\Big)_{t=0} &= \frac{d}{dt}\Big((\mathrm{ad}\, e^{tA}X)Y\Big)_{t=0} + \frac{d}{dt}\Big((\mathrm{ad}\, X)(e^{tA}Y)\Big)_{t=0} \\ &= (\mathrm{ad}\, AX)Y + (\mathrm{ad}\, X)(AY).\end{aligned} \tag{2.36}$$

Finally, $A[X,Y] = [AX,Y] + [X,AY]$.

□

As notational convention, if G and H are Lie groups, their Lie algebra are denoted by $\mathfrak{g}$ and $\mathfrak{h}$.

Lemma 2.18.
Let $\mathfrak{g}$ be a Lie algebra ans $\mathfrak{s}$ be a subset of $\mathfrak{g}$. The algebra of the group generated by $e^{\mathfrak{s}}$ is the algebra generated by $\mathfrak{s}$.

2.5 Lie subgroup

Definition 2.19.
*Let G be a Lie group. A submanifold H of G is a **Lie subgroup** of G when*

(i) as group, H is a subgroup of G,

(ii) H is a topological group.

Remark 2.20.
The definition doesn't include that H has the same topology as G (or the induced one). In some literature, the definition of a Lie subgroup include the fact for H to be a topological subgroup. This choice make some proofs much easier and others more difficult; be careful when you try to compare different texts.

Lemma 2.21.
A Lie subgroup is a Lie group.

(without proof)

Theorem 2.22.
If G is a Lie group, then

(i) if $\mathfrak{h}$ is the Lie algebra of a Lie subgroup H of G, then it is a subalgebra of $\mathfrak{g}$,

(ii) Any subalgebra of $\mathfrak{g}$ is the Lie algebra of one and only one connected Lie subgroup of G.

Problem and misunderstanding 2.
À mon avis, il faut dire "connexe et simplement connexe", et non juste "connexe".

Proof. First item. Let $i: H \to G$ be the identity map; it is a homomorphism from H to G, thus di_e is a homomorphism from $\mathfrak{h}$ to $\mathfrak{g}$. Conclusion: $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$.
Characterization for $\mathfrak{h}$. Before to go on with the second point, we derive an important characterization of $\mathfrak{h}$: formula (2.37). Consider $\exp_H: \mathfrak{h} \to H$ and $\exp_G: \mathfrak{g} \to G$; from unicity of the exponential, for any $X \in \mathfrak{h}$, $\exp_H X = \exp_G X$, so that one can simply write "exp" instead of "$\exp_h$" or "$\exp_G$".

Now, if $X \in \mathfrak{h}$, the map $t \to \exp tX$ is a curve in H. But it is not immediately clear that such a curve in H is automatically build from a vector in $\mathfrak{h}$ rather than in $\mathfrak{g}$. More precisely, consider a $X \in \mathfrak{g}$ such that $t \to \exp tX$ is a path (continuous curve) in H. By lemma 1.4, the map $t \to \exp tX$ is differentiable and thus by derivation, $X \in \mathfrak{h}$. Up to here we had shown that

$$\mathfrak{h} = \{X \in \mathfrak{g} : \text{the map } t \to \exp tX \text{ is a path in } H\}. \tag{2.37}$$

Thus $\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$.
Second item. For the second part, we consider $\mathfrak{h}$ any subalgebra of $\mathfrak{g}$ and H, the smallest subgroup of G which contains $\exp\mathfrak{h}$. We also consider a basis $\{X_1, \ldots, X_n\}$ of $\mathfrak{g}$ such that $\{X_{r+1}, \ldots, X_n\}$ is a basis of $\mathfrak{h}$.

By corollary 3.204, the set of linear combinations of elements of the form $X(M)$ with $M = (0, \ldots, 0, m_{r+1}, \ldots, m_r)$ form a subalgebra of $U(\mathfrak{g})$. If $X = x_1X_1 + \ldots + x_nX_n$, we define $|X| = (x_1^2 + \ldots + x_n^2)^{1/2}$ $(x_i \in \mathbb{R})$.

Let us consider a $\delta > 0$ such that exp is a diffeomorphism (normal neighbourhood) from $B_\delta = \{X \in \mathfrak{g} : |X| < \delta\}$ to a neighbourhood N_e of $e \in G$ and such that $\forall x, y, xy \in N_e$,

$$(xy)_k = \sum_{M,N} C^{[k]}_{MN} x^M y^N \tag{2.38}$$

holds[2]. We note $V = \exp(\mathfrak{h} \cap B_\delta) \subset N_e$. The map

$$\exp(x_{r+1}X_{r+1} + \ldots + x_nX_n) \to (x_{r+1}, \ldots, x_n)$$

is a coordinate system on V for which V is a connected manifold. But $\mathfrak{h} \cap B_\delta$ is a submanifold of B_δ, then V is a submanifold of N_e and consequently of G.

Let x, $y \in V$ such that $xy \in N_e$ (this exist: $x = y = e$); the canonical coordinates of xy are given by (2.38). Since $x_k = y_k = 0$ for $1 \leqslant k \leqslant r$, $(xy)_k = 0$ for the same k because for $(xy)_k$ to be non zero, one need $m_1 = \ldots = m_r = n_1 = \ldots = n_r = 0$ – otherwise, x^M or y^N is zero. Now we looks at $C^{[k]}_{MN}$ for such a k (say $k = 1$ to fix ideas) : $[k] = (\delta_{11}, \ldots, \delta_{1k}) = (1, 0, \ldots, 0)$ and by definition of the C's,

$$X(M)X(N) = \sum_P C^P_{MN} X(P).$$

But we had seen that the set of the $X(A)$ with $A = (0, \ldots, 0, a_{r+1}, \ldots, a_n)$ form a subalgebra of $U(\mathfrak{g})$. Then, only terms with $P = (0, \ldots, 0, p_{r+1}, \ldots, p_n)$ are present in the sum; in particular, $C^{[k]}_{MN} = 0$ for $k = 1, \ldots, r$. Thus $VV \cap N_e \subset V$.

The next step is to consider $\mathcal{V}$, the set of all the subset of H whose contains a neighbourhood of e in V. We can check that this fulfils the six axioms of a topological group :

(i) The intersection of two elements of $\mathcal{V}$ is in $\mathcal{V}$;

(ii) the intersection of all the elements of $\mathcal{V}$ is $\{e\}$;

(iii) any subset of H which contains a set of $\mathcal{V}$ is in $\mathcal{V}$;

(iv) If $\mathcal{U} \in \mathcal{V}$, there exists a $\mathcal{U}_1 \in \mathcal{V}$ such that $\mathcal{U}_1\mathcal{U}_1 \subset \mathcal{U}$ because $VV \cap N_e \subset V$;

(v) if $\mathcal{U} \in \mathcal{V}$, then $\mathcal{U}^{-1} \in \mathcal{V}$ because the inverse map is differentiable and transforms a neighbourhood of e into a neighbourhood of e;

(vi) if $\mathcal{U} \in \mathcal{V}$ and $h \in H$, then $h\mathcal{U}h^{-1} \in \mathcal{V}$.

To see this last item, we denote by log the inverse map of $\exp: B_\delta \to N_e$. By definition of V, it sends V on $\mathfrak{h} \cap B_\delta$. If $X \in \mathfrak{g}$, there exists one and only one $X' \in \mathfrak{g}$ such that $he^{tX}h^{-1} = e^{tX'}$ for any $t \in \mathbb{R}$. Indeed we know that $he^X h^{-1} = e^{\mathrm{Ad}_h X}$, then X' must satisfy $e^{tX'} = e^{\mathrm{Ad}_h tX}$. If it is true for any t, then, by derivation, $X' = \mathrm{Ad}_h X$.

The map $X \to X'$ is an automorphism of $\mathfrak{g}$ which sent $\mathfrak{h}$ on itself. So one can find a δ_1 with $0 < \delta_1 < \delta$ such that

$$h\exp(B_{\delta_1} \cap \mathfrak{h})h^{-1} \subset V.$$

[2]The validity of this second condition is assured during the proof of theorem 3.206 which is not given here.

Indeed, $he^{\mathfrak{h}}h^{-1} \subset \mathfrak{h}$, so that taking $\delta_1 < \delta$, we get the strict inclusion. We can choose δ_1 even smaller to satisfy $he^{B_{\delta_1}}h^{-1} \subset N_e$. Since the map $X \to \log(he^X h^{-1})$ from $B_{\delta_1 \cap \mathfrak{h}}$ to $B_\delta \cap \mathfrak{h}$ is regular, the image of $B_{\delta_1} \cap \mathfrak{h}$ is a neighbourhood of 0 in $\mathfrak{h}$. Thus $he^{B_{\delta_1} \cap \mathfrak{h}}h^{-1}$ is a neighbourhood of e in V. Finally, $h\mathcal{U}h^{-1} \in \mathcal{V}$ and the last axiom of a topological group is checked.

This is important because there exists a topology on H such that H becomes a topological group and $\mathcal{V}$ is a family of neighbourhood of e in H. In particular, V is a neighbourhood of e in H.

For any $z \in G$, we define the map $\phi_z \colon zN_e \to B_\delta$ by

$$\phi_z(ze^{x_1X_1+\ldots+x_nX_n}) = (x_1, \ldots, x_n), \tag{2.39}$$

and we denote by φ_z the restriction of ϕ_z to zV. If $z \in H$, then φ_z sends the neighbourhood zV of z in H to the open set $B_\delta \cap \mathfrak{h}$ in $\mathbb{R}^{n-r}$. Indeed, an element of zV is a ze^Z with $Z \in \mathfrak{h} \cap B_\delta$ which is sent by φ_z to an element of $\mathfrak{h} \cap B_\delta$. (we just have to identify $x_1X_1 + \ldots + x_nX_n$ with $(x_1, \ldots, x_n)$).

Moreover, if $z_1, z_2 \in H$, the map $\varphi_{z_1} \circ \varphi_{z_2}^{-1}$ is the restriction to an open subset of $\mathfrak{h}$ of $\phi_{z_1} \circ \phi_{z_2}$. Then $\varphi_{z_1} \circ \varphi_{z_2}^{-1}$ is differentiable. Conclusion: $(H, \varphi_z : z \in H)$ is a differentiable manifold.

Recall that the definition of $\mathfrak{h}$ was to be a subalgebra of $\mathfrak{g}$; therefore $V = e^{\mathfrak{h} \cap B_\delta}$ is a submanifold of G. But the left translations are diffeomorphism of H and H is the smallest subgroup of G containing $e^{\mathfrak{h}}$. Thus H is a manifold on which the multiplication is diffeomorphic and consequently, H is a Lie subgroup of G.

Rest to prove that the Lie algebra of H is $\mathfrak{h}$ and the unicity part of the theorem.

We know that $\dim H = \dim \mathfrak{h}$ and moreover for $i > r$, the map $t \to \exp tX_i$ is a curve in H. Now, the fact that $\mathfrak{h}$ is the set of $X \in \mathfrak{g}$ such that $t \to \exp tX$ is a path in H show that $X_i \in \mathfrak{h}$. Then the Lie algebra of H is $\mathfrak{h}$ and H is a connected group because it is generated by $\exp \mathfrak{h}$ which is a connected neighbourhood of e in H.

We turn our attention to the unicity part. Let H_1 be a connected Lie subgroup of G such that $T_eH_1 = \mathfrak{h}$. Since $\exp_{\mathfrak{h}} X = \exp_{\mathfrak{h}_1} X$, $H = H_1$ as set. But exp is a differentiable diffeomorphism from a neighbourhood of 0 in $\mathfrak{h}$ to a neighbourhood of e in H and H_1, so as Lie groups, H and H_1 are the same.

□

We state a corollary without proof :

Corollary 2.23.
If H_1 and H_2 are two Lie subgroups of the Lie group G such that $H_1 = H_2$ as topological groups, then $H_1 = H_2$ as Lie groups.

Proposition 2.24.
Let G_1 and G_2 be two Lie groups with same Lie algebra such that $\pi_0(G_1) = \pi_0(G_2)$ and $\pi_1(G_1) = \pi_1(G_2)$, then G_1 and G_2 are isomorphic.

Proof. The assumptions of equality of Lie algebras and of the π_0 make that the universal covering $\tilde{G}_1$ and $\tilde{G}_2$ of G_1 and G_2 are the same. But we know that $G_i = \tilde{G}_i/\pi_1(G_i)$. Now equality $\pi_1(G_1) = \pi_1(G_2)$ concludes that $G_1 = G_2$. □

Lemma 2.25.
Let $\mathfrak{g}$ admit a direct sum decomposition (as vector space) $\mathfrak{g} = \mathfrak{m} \oplus \mathfrak{n}$. Then there exists open and bounded neighbourhoods $\mathcal{U}_m$ and $\mathcal{U}_n$ of 0 in $\mathfrak{m}$ and $\mathfrak{n}$ such that the map

$$\phi \colon \mathcal{U}_m \times \mathcal{U}_n \to G \;\; (A, B) \quad \mapsto e^A e^B \tag{2.40}$$

is a diffeomorphism between $\mathcal{U}_m \times \mathcal{U}_n$ and an open neighbourhood of e in G.

Proof. Let $\{X_1, \ldots, X_n\}$ be a basis of $\mathfrak{g}$ such that $X_i \in \mathfrak{m}$ for $1 \leqslant i \leqslant r$ and $X_j \in \mathfrak{n}$ for $r < j \leqslant n$. We consider $\{t_1, \ldots, t_n\}$, the canonical coordinates of $\exp(x_1X_1 + \ldots + x_rX_r)\exp(x_{r+1}X_{r+1} + \ldots + x_nX_n)$ in this coordinate system. By properties of the exponential, the function φ_j defined by $t_j = \varphi_j(x_1, \ldots, x_n)$ is differentiable at $(0, \ldots, 0)$. If $x_i = \delta_{ij}s$, then $t_i = \delta_{ij}s$ and the Jacobian of

$$\frac{\partial(\varphi_1, \ldots, \varphi_n)}{\partial(x_1, \ldots, x_n)}$$

is 1 for $x_1 = \ldots = x_n = 0$. Thus $d\varphi_e$ is a diffeomorphism and so φ is a locally diffeomorphic. □

Theorem 2.26.
Let G be a Lie group whose Lie algebra is $\mathfrak{g}$ and H, a closed subgroup (not specially a Lie *subgroup) of G. Then there exists one and only one analytic structure on H for which H is a topological Lie subgroup of G.*

Remark 2.27.
A topological Lie subgroup is stronger that a common Lie subgroup because it needs to be a topological subgroup: it must carry exactly *the induced topology. In our definition of a Lie group, this feature doesn't appears.*

Proof. Let $\mathfrak{h}$ be the subspace of $\mathfrak{g}$ defined by

$$\mathfrak{h} = \{X \in \mathfrak{g} \text{ st } \forall t \in \mathbb{R},\, e^{tX} \in H\}. \tag{2.41}$$

We begin to show that $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$; i.e. to show that $t(X+Y) \in \mathfrak{h}$ and $t^2[X,Y] \in \mathfrak{h}$ if X, $Y \in \mathfrak{h}$. Remark that $X \in \mathfrak{h}$ and $s \in \mathbb{R}$ implies $sX \in \mathfrak{h}$. Consider now X, $Y \in \mathfrak{h}$ and the classical formula :

$$\Big(\exp(\frac{t}{n}X)\exp(\frac{t}{n}Y)\Big)^n = \exp(t(X+Y) + \frac{t^2}{2n}[X,Y] + o(\frac{1}{n^2})), \tag{2.42a}$$

$$\Big(\exp(-\frac{t}{n}X)\exp(-\frac{t}{n}Y)\exp(\frac{t}{n}X)\exp(\frac{t}{n}Y)\Big)^{n^2} = \exp\Big(t^2[X,Y] + o(\frac{1}{n})\Big). \tag{2.42b}$$

The left hand side of these equations are in H for any n; but, since H is closed, it keeps in H when $n \to \infty$. The right hand side, at the limit, is just $\exp(t(X+Y))$ and $\exp(t^2[X,Y])$, which keeps in H for any t. Thus $X+Y$ and $[X,Y]$ belong to $\mathfrak{h}$. The space $\mathfrak{h}$ is thus a Lie subalgebra of $\mathfrak{g}$.

Let H^* be the connected Lie subgroup of G whose Lie algebra is $\mathfrak{h}$ (existence and unicity from 2.22). From the proof of theorem 2.22, we know that H^* is the smallest subgroup of G containing $\exp\mathfrak{h}$, then it is made up from products and inverses of elements of the type e^X with $X \in \mathfrak{h}$, and thus is is included in H by definition of $\mathfrak{h}$. So, $H^* \subset H$.

We will show that if we put on H^* the induced topology from G and if H_0 denotes the identity component of H, then $H^* = H_0$ as topological groups. For this, we first have to show the equality as set and then prove that if N is a neighbourhood of e in H^*, then it is a neighbourhood of e in H_0. In facts, the equality as set can be derives from this second fact. Indeed, since H_0 is a connected topological group, it is generated by any neighbourhood of e, so if one can show that any neighbourhood N of e in H^* is a neighbourhood of e in H, then H^* is a neighbourhood of e in H_0 and then H_0 should be generated by H^*, so that $H_0 \subset H^*$ (as set). Moreover, the most general element of H^* is product and inverse of e^X with $X \in \mathfrak{h}$ and e^X is connected to e by the path e^{tX} $(t\colon 1 \to 0)$. Then $H^* \subset H_0$, and $H^* = H_0$ as set. Immediately, $H^* = H_0$ as topological groups from our assertion about neighbourhoods of e. Let us now prove it.

We consider a neighbourhood N of e in H^* and suppose that this is not a neighbourhood of e in H. Thus there exists a sequence $c_k \in H \backslash N$ such that $c_k \to e$ in the sense of the topology on G. Indeed, a neighbourhood of e in the sense of H must contains at least a point which is not in N because if we have an open set of H around e included in N, then N is a neighbourhood of e for H. So we consider a suitable sequence of such open set around e and one element not in N in each of them. There is the c_k's[3].

Using lemma 2.25 with a decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ (i.e. $\mathfrak{m}$: a complementary for $\mathfrak{h}$ for $\mathfrak{g}$), one can find sequences $A_k \in \mathcal{U}_m$ and $B_k \in \mathcal{U}_n$ such that

$$c_k = e^{A_k} e^{B_k}.$$

Here, $\mathcal{U}_m$ is an open neighbourhood of 0 in $\mathfrak{m}$ and $\mathcal{U}_h$, an open neighbourhood of 0 in $\mathfrak{h}$.

As $e^{B_k} \in N$ and $c_k \in H \backslash N$, $A_k \neq 0$ and $\lim A_k = \lim B_k = 0$ (because $(A,B) \to e^A e^B$ is a diffeomorphism and $e^0 e^0 = e$ – and also because all is continuous and thus has a good behaviour with respect to the limit). The set $\mathcal{U}_m$ is open and bounded –this is a part of the lemma. Then there exist a sequence of positive reals numbers $r_k \in$ such that $r_k A_k \in \mathcal{U}_m$ and $(r_k+1)A_k \notin \mathcal{U}_m$. We know that $\mathcal{U}_m$ is a bounded open subset of the vector space $\mathfrak{m}$, then the whole sequences $r_k A_k$ and $(r_k+1)A_k$ are in a compact domain of $\mathfrak{m}$. Then –by eventually considering subsequences– there are no problems to consider limits of these sequences in $\mathfrak{m}$: $r_k A_k \to A \in \mathfrak{m}$ (not necessary in $\mathcal{U}_m$). Since $A_k \to 0$, the point A is the common limit of $r_k A_k \in \mathcal{U}_m$ and of $(r_k+1)A_k \notin \mathcal{U}_m$. Thus A is in the boundary of $\mathcal{U}_m$; in particular, $A \neq 0$.

On the other hand, consider two integers p, q with $q > 0$. One can find sequences $s_k, t_k \in \mathbb{N}$ and $0 \leqslant t_k < q$ such that $pr_k = qs_k + t_k$. It is clear that

$$\lim_{k\to\infty} \frac{t_k}{q} A_k = 0, \tag{2.43}$$

thus

$$\exp\frac{p}{q}A = \lim \exp\frac{pr_k}{a}A_k = \lim(\exp A_k)^{s_k},$$

which belongs to H. By continuity, $\exp tA \in H$ for any $t \in \mathbb{R}$ and finally $A \in \mathfrak{h}$; this contradict $A \neq 0$ so that $A \in \mathfrak{m}$ (because by definition, $A \in \mathfrak{m}$ and the sum $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ is direct).

By its definition, H^* has an analytic structure of Lie subgroup of G; but we had just proved that the induced topology from G is the one of H_0 which by definition is a submanifold of G. So the set $H_0 = H^*$ becomes

[3] Je crois qu'on utilise l'axiome du choix.

a submanifold of G whose topology is compatible with the analytic structure: thus it is a Lie subgroup of G. From analyticity, this structure is extended to the whole H.

Problem and misunderstanding 3.
Est-ce bien vrai, tout ça ? En particulier, je n'utilise pas que H_0 est ouvert dans H (ce qui est un tho de topo classique : je ne vois pas pourquoi Helgason fait tout un cinéma –que je ne comprends pas– dessus). En prenant $N = H^*$, on a juste démontré que H_0 est un voisinage de e dans H, mais ça, on le savait bien avant.

The unicity part comes from the corollary 2.23. □

With the notations and the structure of theorem 2.26, the subgroup H is discrete if and only if $\mathfrak{h} = \{0\}$. Indeed, recall the definition (2.41) :

$$\mathfrak{h} = \{X \in \mathfrak{g} : \forall t \in \mathbb{R}, e^{tX} \in H\},$$

and the fact that there exists a neighbourhood of e in H on which the exponential map is a diffeomorphism.

Remark 2.28.
This fact should not be placed after the following lemma. In fact, we use here just the existence of normal neighbourhood (which is a common result) while the following lemma gives much more than normal neighbourhood.

The lemma (without proof) :

Lemma 2.29.
Let G be a Lie group and H, a Lie subgroup of G ($\mathfrak{g}$ and $\mathfrak{h}$ are the corresponding Lie algebras). If H is a topological subspace of G (cf remark 2.20), then there exists an open neighbourhood V of 0 in $\mathcal{G}$ such that

(i) $\exp$ *is a diffeomorphism between V and an open neighbourhood of e in G,*

(ii) $\exp(V \cap \mathfrak{h}) = (\exp V) \cap H$.

Now a theorem with proof.

Theorem 2.30.
Let G and H be two Lie groups and $\varphi\colon G \to H$ a continuous homomorphism. Then φ is analytic.

Proof. The Lie algebra of the product manifold $G \times H$ as $\mathfrak{g} \times \mathfrak{h}$ is given in 1.11. We define

$$K = \{(g, \varphi(g)) : g \in G\} \subset G \times H. \tag{2.44}$$

It is clear that K is closed in $G \times H$ because G is closed and φ is continuous. By theorem 2.26, there exists an unique differentiable structure on $G \times H$ such that K is a topological Lie subgroup of $G \times H$ (i.e. : Lie subgroup + induced topology). The Lie algebra of K is

$$\mathfrak{k} = \{(X, Y) \in \mathfrak{g} \times \mathfrak{h} : \forall t \in \mathbb{R}, (e^{tX}, e^{tY}) \in K\}. \tag{2.45}$$

Let N_0 be an open neighbourhood of 0 in $\mathfrak{h}$ such that $\exp$ is diffeomorphic between N_0 and an open neighbourhood N_e of e in H. We define M_0 and M_e in the same way, for G instead of H. We can suppose $\varphi(M_e) \subset N_e$: if it is not, we consider a smaller M_e : the openness of N_e and the continuity of φ make it coherent.

The lemma 2.29 allow us to consider M_0 and N_0 small enough to say that

$$\exp\colon (M_0 \times N_0) \cap \mathfrak{k} \to (M_e \times N_e) \cap K$$

is diffeomorphic. Now, we are going to show that for any $X \in \mathfrak{g}$, there exists an unique $Y \in \mathfrak{h}$ such that $(X, Y) \in \mathfrak{k}$. The unicity is easy: consider $(X, Y_1), (X, Y_2) \in \mathfrak{k}$; then $(0, Y_1 - Y_2) \in \mathfrak{k}$ (because a Lie algebra is a vector space). Then the definition of $\mathfrak{k}$ makes for any $t \in \mathbb{R}$, $(e, \exp t(Y_1 - Y_2)) \in K$. Consequently, $\exp t(Y_1 - Y_2) = \varphi(e) = e$ and then $Y_1 - Y_2 = 0$.

In order to show the existence, let us consider a $r > 0$ such that $X_r = (1/r)X$ keeps in M_0. This exists because the sequence $X_r \to 0$ (then it comes M_0 from a certain r). From the definitions, $\exp$ is diffeomorphic between M_0 and M_e, then $\exp X_r \in M_e$ and $\varphi(\exp X_r) \in N_e$ because $\varphi(M_e) \subset N_e$.

From this, there exists an unique $Y_r \in N_0$ such that $\exp Y_r = \varphi(\exp X_r)$ and an unique $Z_r \in (M_0 \times N_0) \cap \mathfrak{k}$ satisfying $\exp Z_r = (\exp X_r, \exp Y_r)$. But $\exp$ is bijective from $M_0 \times N_0$, so that $Z_r = (X_r, Y_r)$ and we can choose $Y = rY_r$ as a $Y \in \mathfrak{h}$ such that $(X, Y) \in \mathfrak{k}$ (it is not really a choice: the unicity was previously shown). We denotes by $\psi\colon \mathfrak{g} \to \mathfrak{h}$ the map which gives the unique $Y \in \mathfrak{h}$ associated with $X \in \mathfrak{g}$ such that $(X, Y) \in \mathfrak{k}$. This is a homomorphism between $\mathfrak{g}$ and $\mathfrak{h}$.

By definition, $(X, \psi(X)) \in \mathfrak{k}$, i.e. $(\exp tX, \exp t\psi(X)) \in K$ or

$$\varphi(\exp tX) = \exp t\psi(X). \tag{2.46}$$

Let us now consider a basis $\{X_1, \ldots, X_n\}$ of $\mathfrak{g}$. Since φ is a homomorphism,

$$\varphi\big((\exp t_1 X_1)(\exp t_2 X_2) \ldots (\exp t_n X_n)\big) = \big(\exp t_1 \psi(X_1)\big) \ldots \big(\exp t_n \psi(X_n)\big) \tag{2.47}$$

Now, we apply lemma 2.25 on the decomposition of $\mathfrak{g}$ into the n subspace spanned by the n vector basis (this is n applications of the lemma), the map

$$(\exp t_1 X_1) \ldots (\exp t_n X_n) \to (t_1, \ldots, t_n)$$

is a coordinate system around e in G. In this case, the relation (2.47) shows that φ is differentiable at e. Then it is differentiable anywhere in G. □

Proposition 2.31.
Let G be a Lie group and H, a Lie subgroup of G ($\mathfrak{g}$ and $\mathfrak{h}$ are the corresponding Lie algebras). We suppose that H has at most a countable number of connected components. Then

$$\mathfrak{h} = \{X \in \mathfrak{g} : \forall t \in \mathbb{R}, e^{tX} \in H\} \tag{2.48}$$

Proof. We will once again use the lemma 2.25 with $\mathfrak{n} = \mathfrak{h}$ and $\mathfrak{m}$, a complementary vector space of $\mathfrak{h}$ in $\mathfrak{g}$. We define

$$V = \exp \mathcal{U}_m \exp \mathcal{U}_h$$

where $\mathcal{U}_m$ and $\mathcal{U}_h$ are the sets given by the lemma. We consider on V the induced topology from G. If we define

$$\mathcal{A} = \{A \in \mathcal{U}_m : e^A \in H\},$$

we have

$$H \cap V = \bigcup_{A \in \mathcal{A}} e^A e^{\mathcal{U}_h}. \tag{2.49}$$

First, the definition of V makes clear that the elements of the form $\exp A \exp \mathcal{U}_h$ are in V. They are also in H because $\exp A \in H$ (definition of $\mathcal{A}$) and $\exp \mathcal{U}_h$ still by definition. In order to see the inverse inclusion, let us consider a $h \in H \cap V$. We know that

$$(A, B) \to \exp A \exp B \tag{2.50}$$

is a diffeomorphism between $\mathcal{U}_m \times \mathcal{U}_h$ and a neighbourhood of e in G which we called V. Thus any element of V (*a fortiori* in $V \cap H$) can be written as $\exp A \exp B$ with $A \in \mathcal{U}_m$ and $B \in \mathcal{U}_h$. Then $h = e^A e^B$ for some $A \in \mathcal{U}_m$, $B \in \mathcal{U}_h$. Since H is a group and $e^B \in H$, in order the product to belongs to H, e^A must lies in H : $A \in \mathcal{A}$.

Remark 2.32.
Note that since (2.50) is diffeomorphic, the union in right hand side of (2.49) is disjoint. Each member of this union is a neighbourhood in H because it is a set $h \exp \mathcal{U}_h$ where $\exp \mathcal{U}_h$ is a neighbourhood of e in H.

Now we consider the map $\pi \colon V \to \mathcal{U}_m$,

$$\pi(e^X e^Y) = X$$

if $X \in \mathcal{U}_m$ and $Y \in \mathcal{U}_h$. This is a continuous map which sends $H \cap V$ into $\mathcal{A}$. The identity component of $H \cap V$ (in the sense of topology of V) is sent to a countable subset of $\mathcal{U}_m$. Indeed by remark 2.32, identity component of $H \cap V$ is only one of the terms in the union (2.49), namely $A = 0$. But we know that $\pi^{-1}(o) = \exp \mathcal{U}_h$, thus $\exp \mathcal{U}_h$ is the identity component of $H \cap V$ for the topology of V.

Let us consider a $X \in \mathfrak{g}$ such that $\exp tX \in H$ for any $t \in \mathbb{R}$, and the map $\varphi \colon \mathbb{R} \to G$, $\varphi(t) = \exp tX$. This is continuous, then there exists a connected neighbourhood $\mathcal{U}$ of 0 in $\mathbb{R}$ such that $\varphi(\mathcal{U}) \subset V$. Then $\varphi(\mathcal{U}) \subset H \cap V$ and the connectedness of $\varphi(\mathcal{U})$ makes $\varphi(\mathcal{U}) \subset \exp \mathcal{U}_h$. But $\exp \mathcal{U}_h$ is an arbitrary small neighbourhood of e in H; the conclusion is that φ is a continuous map from $\mathbb{R}$ into H. Indeed, we had chosen X such that $\exp tX \in H$.

Moreover, we know that

$$e^{(t_0+\epsilon)X} = e^{t_0 X} e^{\epsilon X},$$

but $\exp \epsilon X$ can be as close to e as we want (this proves the continuity at t_0). Then φ is a path in H.

In definitive, we had shown that $\exp tX \in H$ implies that $t \to \exp tX$ is a path. Now equation (2.37) gives the thesis.

□

Corollary 2.33.
Let G be a Lie group and H_1, H_2, two subgroups both having a finite number of connected components (each for his own topology). If $H_1 = H_2$ as sets, then $H_1 = H_2$ as Lie groups.

Proof. The proposition shows that H_1 and H_2 have same Lie algebra. But any Lie subalgebra of $\mathfrak{g}$ is the Lie algebra of exactly one connected subgroup of G (theorem 2.22). Then as Lie groups, $H_{1_0} = H_{2_0}$. Since H_1 and H_2 are topological groups, the equality of they topology on one connected component gives the equality everywhere (because translations are differentiable). □

Definition 2.34.
*A **differentiable subgroup** is a connected Lie subgroup.*

Corollary 2.35.
Let G be a Lie group, and K, H two differentiable subgroups of G. We suppose $K \subset H$. Then K is a differentiable subgroup of the Lie group H.

Proof. The Lie algebras of K and H are respectively denoted by $\mathfrak{k}$ and $\mathfrak{h}$. We denote by K^* the differentiable subgroup of H which has $\mathfrak{k}$ as Lie algebra. The differentiable subgroups K and K^* have same Lie algebra, and then coincide as Lie groups. □

Consider the group $T = S^1 \times S^1$ and the continuous map $\gamma\colon \mathbb{R} \to T$ given by

$$\gamma(t) = (e^{it}, e^{i\alpha t})$$

with a certain irrational α in such a manner that γ is injective and $\Gamma = \gamma(\mathbb{R})$ is dense in T.

The subset Γ is not closed because his complementary in T is not open: any neighbourhood of element $p \in T$ which don't lie in Γ contains some elements of Γ. We will show that the inclusion map $\iota\colon \Gamma \to T$ is continuous. An open subset of T is somethings like

$$\mathcal{O} = (e^{iU}, e^{iV})$$

where U, V are open subsets of $\mathbb{R}$. It is clear that

$$\iota^{-1}(\mathcal{O}) = \{\gamma(t) \text{ st } t \in U + 2k\pi, \alpha t \in V + 2m\pi\},$$

but the set of elements t of $\mathbb{R}$ which satisfies it is clearly open. Then Γ has at least the induced topology from T (as shown in proposition 1.3). In fact, the own topology of Γ is *more* than the induced: the open subsets of Γ whose are just some small segments clearly doesn't appear in the induced topology. Thus the present case is an example (and not a counter-example) of the next theorem at page 36.

This example show the importance of the condition for a topological subspace to have *exactly* the induced topology. If not, any Lie subgroup were a topological Lie subgroup because a submanifold has at least the induced topology. We will go further with this example after the proof.

Theorem 2.36.
Let G be a Lie group and H, a Lie subgroup of G.

(i) If H is a topological Lie subgroup of G, then it is closed in G,

(ii) If H has at most a countable number of connected components and is closed in G, then H is a topological subgroup of G.

Proof. First point. It is sufficient to prove that if a sequence $h_n \in H$ converges (in G) to $g \in G$, then $g \in H$ (this is almost the definition of a closed subset). We consider V, a neighbourhood of 0 in $\mathfrak{g}$ such that

- exp is diffeomorphic between V and an open neighbourhood of e in G,
- $\exp(V \cap \mathfrak{h}) = (\exp V) \cap H$.

This exists by the lemma 2.29; we can suppose that V is bounded. Consider $\mathcal{U}$, an open neighbourhood of 0 in $\mathfrak{g}$ contained in V such that $\exp -\mathcal{U} \exp\mathcal{U} \subset \exp V$.

Since $h_n \to g$, there exists a $N \in \mathbb{N}$ such that $n \geqslant N$ implies $h_n \in g\exp\mathcal{U}$ (i.e h_n is the product of g by an element rather close to e; since the multiplication is differentiable, the notion of "not so far" is good to express the convergence notion). From now we only consider such elements in the sequence. So, $h_N^{-1}h_n \in (\exp V) \cap H$ ($n \geqslant N$) because

$$h_N^{-1}h_n \in \exp -\mathcal{U}g^{-1}g\exp\mathcal{U} \subset \exp V.$$

(note that H is a group, then $h_i^{-1} \in H$) From the second point of the definition of V, there exists a $X_n \in V \cap \mathfrak{h}$ such that $h_N^{-1} h_n = \exp X_n$ for any $n \geqslant N$.

Since V in bounded, there exists a subsequence out of (X_i) (which is also called X_i) converging to a certain $Z \in \mathfrak{g}$. But $\mathfrak{h}$ is closed in $\mathfrak{g}$ because it is a vector subspace (we are in a finite dimensional case), then $Z \in \mathfrak{h}$ and thus the sequence (h_i) converges to $h_N \exp Z$; therefore $g \in H$.

Second point. The subgroup H is closed in G and has a countable number of connected component. Since H is closed, theorem 2.26 it has an analytics structure for which it is a topological Lie subgroup of G. We denotes by H' this Lie group.

The identity map $I \colon H \to H'$ is continuous[4] (see error **??**). Thus any connected component of H is contained in a connected component of H', the it has only a countable number of connected components. By corollary 2.23, $H = H'$ as Lie group.

□

Now we take back our example with $G = S^1 \times S^1$, $H = \gamma(\mathbb{R})$. In this case, the theorem doesn't works. Let us see why as deep as possible. We have $\mathfrak{g} = \mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2$ and $\mathfrak{h} = \mathbb{R}$, a one-dimensional vector subspace of $\mathfrak{g}$. ($\mathfrak{h}$ is a "direction " in $\mathfrak{g}$) First, we build the neighbourhood V of 0 in $\mathfrak{g}$. It is standard to require that exp is diffeomorphic between V and an open around $(1,1) \in S^1 \times S^1$. It also must satisfy $e^{V \cap \mathfrak{h}} = e^V \cap H$. This second requirement is impossible.

Intuitively. We can see $V \subset \mathfrak{g}$ as a little disk tangent to the torus. The exponential map deposits it on the torus, as well that e^V covers a little area on G. Then $e^V \cap H$ is one of these amazing open subset of Γ which are dense in a certain domain of G.

On the other hand, $V \cap \mathfrak{h}$ is just a little vector in $\mathfrak{h}$; the exponential deposits it on a small line in G. This is not the same at all. Then lemma 2.29 fails in our case. Let us review the proof of this lemma until we find a problem.

Let $W_0 \subset \mathfrak{g}$ be a neighbourhood of 0 which is in bijection with an open around e in G. We consider N_0, an open subset of H such that $N_0 \subset W_0$ and N_0 is in bijection with N_e, a neighbourhood of e in G. Until here, no problems. But now the proof says that there exists an open U_e in G such that $N_e = U_e \cap H$. This is false in our case. Indeed, $N_e = e^{N_0}$ is just a segment in G while any subset of G of the form $U_e \cap H$ is an "amazing" open.

So we see that deeply, the obstruction for a Lie subgroup to be a topological Lie subgroup resides in the fact that the topology of a submanifold is *more* than the induced topology, so that we can't automatically find the open U_e in G.

[4] pourquoi ?

Chapter 3

Lie algebras

Sources [8, 9, 11–14].

Definition 3.1.
*A **Lie algebra** is a vector space $\mathfrak{g}$ on $\mathbb{K}(=\mathbb{R},\mathbb{C})$ endowed with a bilinear operation $(x,y)\mapsto[x,y]$ from $\mathfrak{g}\times\mathfrak{g}$ with the properties*

(i) $[x,y]=-[y,x]$

(ii) $\big[x,[y,z]\big]+\big[y,[z,x]\big]+\big[z,[x,y]\big]=0.$

*The second condition is the **Jacobi identity**.*

3.1 Adjoint group

Let $\mathfrak{a}$ be a *real* Lie algebra. We denote by $GL(\mathfrak{a})$ the group of all the nonsingular endomorphism of $\mathfrak{a}$: the linear and nondegenerate operators on $\mathfrak{a}$ as vector space. An element $\sigma\in\mathrm{GL}(\mathfrak{a})$ does not specially fulfils somethings like $\sigma[X,Y]=[\sigma X,\sigma Y]$. The Lie algebra $\mathfrak{gl}(\mathfrak{a})$ is the vector space of the endomorphism (without non degeneracy condition) endowed with the usual bracket $(\operatorname{ad}A)B=[A,B]=A\circ B-B\circ A$. The map $X\to\operatorname{ad}X$ is a homomorphism from $\mathfrak{a}$ to the subalgebra $\operatorname{ad}(\mathfrak{a})$ of $\mathfrak{gl}(\mathfrak{a})$.

The group $\mathrm{Int}(\mathfrak{a})$ is the analytic Lie subgroup of $\mathrm{GL}(\mathfrak{a})$ whose Lie algebra is $\operatorname{ad}(\mathfrak{a})$ by theorem 2.22. This is the **adjoint group** of $\mathfrak{a}$.

Proposition 3.2.
The group $\mathrm{Aut}(\mathfrak{a})$ of all the automorphism of $\mathfrak{a}$ is a closed subgroup of $\mathrm{GL}(\mathfrak{a})$.

Proof. The property which distinguish the elements in $\mathrm{Aut}(\mathfrak{a})$ from the "commons" elements of $\mathrm{GL}(\mathfrak{a})$ is the preserving of structure: $\varphi[A,B]=[\varphi A,\varphi B]$. These are equalities, and we know that a subset of a manifold which is given by some equalities is closed. □

Now, theorem 2.26 provides us an unique analytic structure on $\mathrm{Aut}(\mathfrak{a})$ in which it is a topological Lie subgroup of $\mathrm{GL}(\mathfrak{a})$. From now we only consider this structure. We denote by $\partial(\mathfrak{a})$ the Lie algebra of $\mathrm{Aut}(\mathfrak{a})$: this is the set of the endomorphism D of $\mathfrak{a}$ such that $\forall t\in\mathbb{R}$, $e^{tD}\in\mathrm{Aut}(\mathfrak{a})$. By differencing the equality

$$e^{tD}[X,Y]=[e^{tD}X,e^{tD}Y] \tag{3.1}$$

with respect to t, we see[1] that D is a **derivation** of $\mathfrak{a}$:

$$D[X,Y]=[DX,Y]+[X,DY] \tag{3.2}$$

for any X, $Y\in\mathfrak{a}$. Conversely, consider D, any derivation of $\mathfrak{a}$; by induction,

$$D^k[X,Y]=\sum_{i+j=k}\frac{k!}{i!j!}[D^iX,D^jY] \tag{3.3}$$

[1] As usual, if we consider a basis of $\mathfrak{a}$ as vector space, the expression in the right hand side of

$$[e^{tD}X,e^{tD}Y]=\operatorname{ad}(e^{tD}X)e^{tD}X$$

can be seen as a product matrix times vector, so that Leibnitz works.

where by convention, D^0 is the identity in $\mathfrak{a}$. This relation shows that D fulfils condition (3.1), so that any derivation of $\mathfrak{a}$ lies in $\partial(\mathfrak{a})$. Then

$$\partial(\mathfrak{a}) = \{\text{derivations of } \mathfrak{a}\}.$$

The Jacobi identities show that

$$\mathrm{ad}(\mathfrak{a}) \subset \partial(\mathfrak{a}).$$

From this, we deduce :

$$\mathrm{Int}(\mathfrak{a}) \subset \mathrm{Aut}(\mathfrak{a}). \tag{3.4}$$

(cf. error **??**) Indeed the group $\mathrm{Int}(\mathfrak{a})$ being connected, it is generated[2] by any neighbourhood of e; note that $\mathrm{Aut}(\mathfrak{a})$ has not specially this property. We take a neighbourhood of e in $\mathrm{Int}(\mathfrak{a})$ under the form $\exp V$ where V is a sufficiently small neighbourhood of 0 in $\mathrm{ad}(\mathfrak{a})$ to be a neighbourhood of 0 in $\partial(\mathfrak{a})$ on which exp is a diffeomorphism. In this case, $\exp V \subset \mathrm{Aut}(\mathfrak{a})$ and then $\mathrm{Int}(\mathfrak{a}) \subset \mathrm{Aut}(\mathfrak{a})$.

Elements of $\mathrm{ad}(\mathfrak{a})$ are the **inner derivations** while the ones of $\mathrm{Int}(\mathfrak{a})$ are the **inner automorphism.**

Let $\mathcal{O}$ be an open subset of $\mathrm{Aut}(\mathfrak{a})$; for a certain open subset U of $\mathrm{GL}(\mathfrak{a})$, $\mathcal{O} = U \cap \mathrm{Aut}(\mathfrak{a})$. Then

$$\iota^{-1}(\mathcal{O}) = \mathcal{O} \cap \mathrm{Int}(\mathfrak{a}) = U \cap \mathrm{Aut}(\mathfrak{a}) \cap \mathrm{Int}(\mathfrak{a}) = U \cap \mathrm{Int}(\mathfrak{a}). \tag{3.5}$$

The subset $U \cap \mathrm{Int}(\mathfrak{a})$ is open in $\mathrm{Int}(\mathfrak{a})$ for the topology because $\mathrm{Int}(\mathfrak{a})$ is a Lie[3] subgroup of $\mathrm{GL}(\mathfrak{a})$ and thus has at least the induced topology. This proves that the inclusion map $\iota\colon \mathrm{Int}(\mathfrak{a}) \to \mathrm{Aut}(\mathfrak{a})$ is continuous.

The lemma 1.4 and the consequence below makes $\mathrm{Int}(\mathfrak{a})$ a Lie subgroup of $\mathrm{Aut}(\mathfrak{a})$. Indeed $\mathrm{Int}(\mathfrak{a})$ and $\mathrm{Aut}(\mathfrak{a})$ are both submanifolds of $\mathrm{GL}(\mathfrak{a})$ which satisfy (3.4). By definition, $\mathrm{Aut}(\mathfrak{a})$ has the induced topology from $\mathrm{GL}(\mathfrak{a})$. Then $\mathrm{Int}(\mathfrak{a})$ is a submanifold of $\mathrm{Aut}(\mathfrak{a})$. This is also a subgroup and a topological group ($\mathrm{Int}(\mathfrak{a})$ is not a topological subgroup of $\mathrm{Aut}(\mathfrak{a})$, cf remark 2.20). Then $\mathrm{Int}(\mathfrak{a})$ is a Lie subgroup of $\mathrm{Aut}(\mathfrak{a})$.

Schematically, links between $\mathrm{Int}\,\mathfrak{g}$, $\mathrm{ad}\,\mathfrak{g}$, $\mathrm{Aut}\,\mathfrak{g}$ and $\partial\mathfrak{g}$ are

$$\mathrm{Int}\,\mathfrak{g} \longleftarrow \mathrm{ad}\,\mathfrak{g} \tag{3.6a}$$
$$\mathrm{Aut}\,\mathfrak{g} \longrightarrow \partial\mathfrak{g}. \tag{3.6b}$$

Remark that the sense of the arrows is important. By definition $\partial\mathfrak{g}$ is the Lie algebra of $\mathrm{Aut}\,\mathfrak{g}$, then there exist some algebras $\mathfrak{g}$ and $\mathfrak{g}'$ with $\mathrm{Aut}\,\mathfrak{g} \neq \mathrm{Aut}\,\mathfrak{g}'$ but with $\partial\mathfrak{g} = \partial\mathfrak{g}'$, because the equality of two Lie algebras doesn't implies the equality of the groups. The case of $\mathrm{Int}\,\mathfrak{g}$ and $\mathrm{ad}\,\mathfrak{g}$ is very different: the group is defined from the algebra, so that $\mathrm{ad}\,\mathfrak{g} = \mathrm{ad}\,\mathfrak{g}'$ implies $\mathrm{Int}\,\mathfrak{g} = \mathrm{Int}\,\mathfrak{g}'$ and $\mathrm{Int}\,\mathfrak{g} = \mathrm{Int}\,\mathfrak{g}'$ if and only if $\mathrm{ad}\,\mathfrak{g} = \mathrm{ad}\,\mathfrak{g}'$.

Proposition 3.3.
The group $\mathrm{Int}(\mathfrak{a})$ *is a normal subgroup of* $\mathrm{Aut}(\mathfrak{a})$.

Proof. Let us consider a $s \in \mathrm{Aut}(\mathfrak{a})$. The map $\sigma_s\colon \mathrm{Aut}(\mathfrak{a}) \to \mathrm{Aut}(\mathfrak{a})$, $\sigma_s(g) = sgs^{-1}$ is an automorphism of $\mathrm{Aut}(\mathfrak{a})$. Indeed, consider $g, h \in \mathrm{Aut}(\mathfrak{a})$; direct computations show that $\sigma_s(gh) = \sigma_s(g)\sigma_s(h)$ and $[\sigma_s(g), \sigma_s(h)] = \sigma_s([g,h])$. From this, $(d\sigma_s)_e$ is an automorphism of $\partial(\mathfrak{a})$, the Lie algebra of $\mathrm{Aut}(\mathfrak{a})$. For any $D \in \partial(\mathfrak{a})$ we have

$$(d\sigma_s)_e D = \frac{d}{dt}\Big[sD(t)s^{-1}\Big]_{t=0} = sDs^{-1}. \tag{3.7}$$

Since s is an automorphism of $\mathfrak{a}$ and $\mathrm{ad}(\mathfrak{a})$, a subalgebra of $\mathfrak{gl}(\mathfrak{a})$,

$$s\,\mathrm{ad}\,Xs^{-1} = \mathrm{ad}(sX) \tag{3.8}$$

for any $X \in \mathfrak{a}$, $s \in \mathrm{Aut}(\mathfrak{a})$. Since $\mathrm{ad}(\mathfrak{a}) \subset \partial(\mathfrak{a})$, we can write (3.7) with $D = \mathrm{ad}\,X$ and put it in (3.8) :

$$(d\sigma)_e\,\mathrm{ad}\,X = s\,\mathrm{ad}\,Xs^{-1} = \mathrm{ad}(s\cdot X).$$

We know from general theory of linear operators on vector spaces that if A, B are endomorphism of a vector space and if A^{-1} exists, then $Ae^BA^{-1} = e^{ABA^{-1}}$. We write it with $A = s$ and $B = \mathrm{ad}\,X$:

$$\sigma_s \cdot e^{\mathrm{ad}\,X} = se^{\mathrm{ad}\,X}s^{-1} = e^{s\,\mathrm{ad}\,Xs^{-1}} = e^{\mathrm{ad}(s\cdot X)},$$

sot that

$$\sigma_s \cdot e^{\mathrm{ad}\,X} = e^{\mathrm{ad}(sX)}. \tag{3.9}$$

Ont the other hand, we know that $\mathrm{Int}(\mathfrak{a})$ is connected, so it is generated by elements of the form $e^{\mathrm{ad}\,X}$ for $X \in \mathfrak{a}$. Then $\mathrm{Int}(\mathfrak{a})$ is a normal subgroup of $\mathrm{Aut}(\mathfrak{a})$; the automorphism s of $\mathfrak{a}$ induces the isomorphism $g \to sgs^{-1}$ in $\mathrm{Int}(\mathfrak{a})$ because of equation (3.9). □

[2] See proposition 2.1
[3] Is it true ??

More generally, if s is an isomorphism from a Lie algebra $\mathfrak{a}$ to a Lie algebra $\mathfrak{b}$, then the map $g \to sgs^{-1}$ is an isomorphism between $\mathrm{Aut}(\mathfrak{a})$ and $\mathrm{Aut}(\mathfrak{b})$ which sends $\mathrm{Int}(\mathfrak{a})$ to $\mathrm{Int}(\mathfrak{b})$. Indeed, consider an isomorphism $s\colon \mathfrak{a} \to \mathfrak{b}$ and $g \in \mathrm{Aut}(\mathfrak{a})$. If $g \in \mathrm{Int}(\mathfrak{a})$, we have to see that $sgs^{-1} \in \mathrm{Int}(\mathfrak{b})$. By definition, $\mathrm{Int}(\mathfrak{a})$ is the analytic subgroup of $\mathrm{GL}(\mathfrak{a})$ which has $\mathrm{ad}(\mathfrak{a})$ as Lie algebra. We have $g = e^{\mathrm{ad}\, A}$, then $sgs^{-1} = e^{\mathrm{ad}(sA)}$ which lies well in $\mathrm{Int}(\mathfrak{b})$.

Lemma 3.4.
The adjoint map is an homomorphism $\mathrm{ad}\colon \mathfrak{g} \to \mathrm{GL}(\mathfrak{g})$. *In other terms for every* $X, Y \in \mathfrak{g}$ *we have*

$$\big[\mathrm{ad}(X), \mathrm{ad}(Y)\big] = \mathrm{ad}\,\big([X,Y]\big) \tag{3.10}$$

as operators on $\mathfrak{g}$. *In particular the algebra acts on itself and* $\mathfrak{g}$ *carries a representation of each of its subalgebra.*

Proof. Using the fact that $\mathrm{ad}(X)$ is a derivation and Jacobi, for $Z \in \mathfrak{g}$ we have

$$\big[\mathrm{ad}(X), \mathrm{ad}(Y)\big]Z = \mathrm{ad}(X)\,\mathrm{ad}(Y)Z - \mathrm{ad}(Y)\,\mathrm{ad}(X)Z \tag{3.11a}$$
$$= \big[[X,Y],Z\big] + \big[Y,[X,Z]\big] - \big[[Y,X],Z\big] - \big[X,[Y,Z]\big] \tag{3.11b}$$
$$= \mathrm{ad}\,\big([X,Y]\big)Z. \tag{3.11c}$$

□

3.2 Adjoint representation

Let G be a Lie group and $g \in G$; one can consider the map $I\colon G \times G \to G$ given by $I(g)h = ghg^{-1}$. Seen as $I(g)\colon G \to G$, this is an analytic automorphism of G. We define :

$$\mathrm{Ad}(g) = dI(g)_e.$$

Using equation $\varphi(\exp X) = \exp d\varphi_e(X)$ with $\varphi = I(g)$,

$$ge^X g^{-1} = \exp[\mathrm{Ad}(g)X] \tag{3.12}$$

for every $g \in G$ and $X \in \mathfrak{g}$. The map $g \to \mathrm{Ad}(g)$ is a homomorphism from G to $\mathrm{GL}(\mathfrak{g})$. This homomorphism is called the **adjoint representation** of G.

Proposition 3.5.
The adjoint representation is analytic.

Proof. We have to prove that for any $X \in \mathfrak{g}$ and for any linear map $\omega\colon \mathfrak{g} \to \mathbb{R}$, the function $\omega(\mathrm{Ad}(g)X)$ is analytic at $g = e$. Indeed if we take as ω , the projection to the ith component and X as the jth basis vector ($\mathfrak{g}$ seen as a vector space), and if we see the product $\mathrm{Ad}(g)X$ as a product matrix times vector, $(\mathrm{Ad}(g)X)_i$ is just $\mathrm{Ad}(g)_{ij}$. Then our supposition is the analyticity of $g \to \mathrm{Ad}(g)_{ij}$ at $g = e$. [4]

Now we prove it. Consider $f \in C^\infty(G)$, analytic at $g = e$ and such that $Yf = \omega(Y)$ for any $Y \in \mathfrak{g}$. Using equation (3.12),

$$\omega(\mathrm{Ad}(g)X) = (\mathrm{Ad}(g)X)f = \frac{d}{dt}\Big[f(e^{t\,\mathrm{Ad}(g)X})\Big]_{t=0} = \frac{d}{dt}\Big[f(ge^{tX}g^{-1})\Big]_{t=0}, \tag{3.13}$$

which is well analytic at $g = e$. □

Proposition 3.6.
Let G be a connected Lie group and H, an analytic subgroup of G. Then H is a normal subgroup of G if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

Proof. We consider X, $Y \in \mathfrak{g}$. Formula $\exp tX \exp tY \exp -tY = \exp(tY + t^2[X,Y] + o(t^3))$ and equation (3.12) give

$$\exp\Big(\mathrm{Ad}(e^{tX})tY\Big) = \exp\Big(tY + t^2[X,Y] + o(t^3)\Big).$$

Since it is true for any X, $Y \in \mathfrak{g}$, $\mathrm{Ad}(e^{tX})tY = tY + t^2[X,Y]$; thus

$$\mathrm{Ad}(e^{tX}) = \mathbb{1} + t[X,Y] + o(t^2). \tag{3.14}$$

[4] L'analicité de Ad, elle vient par prolongement analytique depuis juste un point ?

Since we know that $d\,\mathrm{Ad}_e\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$ is a homomorphism (Ad is seen as a map $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$), taking the derivative of the last equation with respect to t gives

$$d\,\mathrm{Ad}_e(X) = \mathrm{ad}\,X. \tag{3.15}$$

Then $\mathrm{Ad}(e^X) = e^{\mathrm{ad}\,X}$. Since is connected, an element of G can be written as $\exp X$ for a certain $X \in \mathfrak{g}$[5]. The purpose is to prove that $g \exp X g^{-1} = \exp(\mathrm{Ad}(g)X)$ remains in H for any $g \in G$ if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. In other words, we want $\mathrm{Ad}(g)X \in \mathfrak{h}$ if and only if $\mathfrak{h}$ is an ideal. We can write $g = e^Y$ for a certain $Y \in \mathfrak{g}$. Thus

$$\mathrm{Ad}(g)X = \mathrm{Ad}(e^Y)X = e^{\mathrm{ad}\,Y}X.$$

Using the expansion

$$e^{\mathrm{ad}\,Y} = \sum_k \frac{1}{k!}(\mathrm{ad}\,Y)^k, \tag{3.16}$$

we have the thesis.

□

Lemma 3.7.
Let G be a connected Lie group with Lie algebra $\mathfrak{g}$. If $\varphi\colon G \to X$ is an analytic homomorphism (X is a Lie group with Lie algebra $\mathfrak{x}$), then

(i) The kernel $\varphi^{-1}(e)$ is a topological Lie subgroup of G; his algebra is the kernel of $d\varphi_e$.

(ii) The image $\varphi(G)$ is a Lie subgroup of X whose Lie algebra is $d\varphi(\mathfrak{g}) \subset \mathfrak{x}$.

(iii) The quotient group $G/\varphi^{-1}(e)$ with his canonical analytic structure is a Lie group. The map $g\varphi^{-1}(e) \mapsto \varphi(g)$ is an analytic isomorphism $G/\varphi^{-1}(e) \to \varphi(G)$. In particular the map $\varphi\colon G \to \varphi(G)$ is analytic.

Proof. First item. We know that a subgroup H closed in G admits an unique analytic structure such that H becomes a topological Lie subgroup of G. This is the case of $\varphi^{-1}(e)$. We know that $Z \in \mathfrak{g}$ belongs to the Lie algebra of $\varphi^{-1}(e)$ if and only if $\varphi(\exp tZ) = e$ for any $t \in \mathbb{R}$. But $\varphi(\exp tZ) = \exp(td\varphi(Z)) = e$ if and only if $d\varphi(Z) = 0$.
Second item. Consider X_1, the analytic subgroup of X whose Lie algebra is $d\varphi(\mathfrak{g})$. The group $\varphi(G)$ is generated by the elements of the form $\varphi(\exp Z)$ for $Z \in \mathfrak{g}$. The group X_1 is generated by the $\exp(d\varphi Z)$. Because of lemma 2.13, these two are the same. Then $\varphi(G) = X_1$ and their Lie algebras are the same.
Third item. We consider H, a closed normal subgroup of G; this is a topological subgroup and the quotient G/H has an unique analytic structure such that the map $G \times G/H \to G/H$, $(g,[x]) \to [gx]$ is analytic. We consider a decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ and we looks at the restriction $\psi\colon \mathfrak{m} \to G$ of the exponential. Then there exists a neighbourhood U of 0 in $\mathfrak{m}$ which is homomorphically send by ψ into an open neighbourhood of e in G and such that $\pi\colon G \to G/H$ sends homomorphically $\psi(U)$ to a neighbourhood of $p_0 \in G/H$ (cf. lemma **??**).

We consider $\mathring{U}$, the interior of U and $B = \psi(\mathring{U})$. The following diagram is commutative :

$$\begin{tikzcd} G \times G/H \arrow[rr, "\Phi"] \arrow[dr, "\pi \times I"'] & & G/H \\ & G/H \times G/H \arrow[ur, "\alpha"'] & \end{tikzcd} \tag{3.17}$$

with $\Phi(g,[x]) = [g^{-1}x]$, $(\pi \times I)(g,[x]) = ([g],[x])$ and $\alpha([g],[x]) = [g^{-1}x]$. Indeed,

$$\alpha \circ (\pi \times I)(g,[x]) = \alpha([g],[x]) = [g^{-1}x].$$

In order to see that α is well defined, remark that if $[h] = [g]$ and $[y] = [x]$ $[g^{-1}x] = [h^{-1}y]$ because H is a normal subgroup of G.

Now, we consider $g_0, x_0 \in G$ and the restriction of $(\pi \times I)$ to $(g_0 B) \times (G/H)$. Since π is homeomorphic on $\psi(U)$ and $B = \psi(\mathring{U})$, on $g_0 B$, π is a diffeomorphism (because the multiplication is diffeomorphic as well)

Problem and misunderstanding 4.
Why is the π a diffeomorphism ? I understand why it is qn homeomorphism, but no more. This is related to problem **??**.

[5] Because G is generated by any neighbourhood of e and there exists such a neighbourhood of e which is diffeomorphic to a subset of $\mathfrak{g}$ by exp.

This diffeomorphism maps to a neighbourhood N of $([g_0],[x_0])$ in $G/H \times G/H$. From the commutativity, we know that $\alpha = \Phi \circ (\pi \times I)^{-1}$, so that α is analytic. Consequently, G/H is a Lie group. On N, α is analytic, then $\alpha(N)$ is analytic.

All this is for a closed normal subgroup H of G. Now we consider $H = \varphi^{-1}(e)$ and $\mathfrak{h}$, the Lie algebra of H. From the first item, we know that the Lie algebra of H is the kernel of $d\varphi : \mathfrak{h} = d\varphi^{-1}(0)$ which is an ideal in $\mathfrak{g}$.

From the second point, the Lie algebra of G/H is $d\pi(\mathfrak{g})$ which is isomorphic to $\mathfrak{g}/\mathfrak{h}$; the bijection is $\gamma(d\pi(X)) = [X] \in \mathfrak{g}/\mathfrak{h}$. In order to prove the injectivity, let us consider $\gamma(A) = \gamma(B)$; $A = d\pi(X)$, $B = d\pi(Y)$. The condition is $[X] = [Y]$; thus it is clear that $d\pi(X) = d\pi(Y)$

Let us consider on the other hand the map $Z + \mathfrak{h} \to d\varphi(Z)$ for $Z \in \mathfrak{g}$[6]. In other words, the map is $[Z] \to d\varphi(Z)$. This is an isomorphism $\mathfrak{g}/\mathfrak{h} \to d\varphi(\mathfrak{g})$, which gives a local isomorphism between G/H and $\varphi(G)$. This local isomorphism is $[g] \to \varphi(g)$ for g in a certain neighbourhood of e in G.

Since $[g] \to \varphi(g)$ has a differential which is an isomorphism, this is analytic at e. Then it is analytic everywhere.

□

Corollary 3.8.
If G is a connected Lie group and if Z is the center of G, then

(i) Ad_G *is an analytic homomorphism from G to* $\mathrm{Int}(G)$*, with kernel Z,*

(ii) the map $[g] \to \mathrm{Ad}_G(g)$ *is an analytic isomorphism from G/Z to* $\mathrm{Int}(\mathfrak{g})$ *(the class $[g]$ is taken with respect to Z).*

Proof. First item. A connected Lie group is generated by a neighbourhood of identity, and any element of a suitable such neighbourhood can be written as the exponential of an element in the Lie algebra. So $\mathrm{Int}(\mathfrak{g})$ is generated by elements of the form $\exp(\mathrm{ad}\, X) = \mathrm{Ad}(\exp X)$; this shows that $\mathrm{Int}(\mathfrak{g}) \subset \mathrm{Ad}(G)$. In order to find the kernel, we have to see $\mathrm{Ad}_G^{-1}(e)$ by the formula

$$e^{\mathrm{Ad}(g)X} = g e^X g^{-1}.$$

We have to find the $g \in G$ such that $\forall X \in \mathfrak{g}$, $\mathrm{Ad}_G(g)X = X$. We taking the exponential of the two sides and using (3.12),

$$g e^X g^{-1} = e^X. \tag{3.18}$$

Then g must commute with any $e^X \in G$: in other words, g is in the kernel of G.
Second item. This is contained in lemma 3.7. Indeed G is connected and we had just proved that $\mathrm{Ad}_G \colon G \to \mathrm{Int}(\mathfrak{g})$ with kernel Z; the third item of lemma 3.7 makes G/Z a Lie group and the map $[g] \to \mathrm{Ad}_G(g)$ an analytic isomorphism from G/Z to $\mathrm{Ad}_G(G) = \mathrm{Int}(\mathfrak{g})$. □

Lemma 3.9.
Let G_1 and G_2 be two locally isomorphic connected Lie groups with trivial center (i.e. $\mathfrak{g}_1 = \mathfrak{g}_2 = \mathfrak{g}$ and $Z(G_i) = \{e\}$). In this case, we have $G_1 = G_2 = \mathrm{Int}(\mathfrak{g})$ where $\mathrm{Int}\,\mathfrak{g}$ *stands for the group of internal automorphism of $\mathfrak{g}$.*

Proof. We denote by G_0 the group $\mathrm{Int}\,\mathfrak{g}$. The adjoint actions $\mathrm{Ad}_i \colon G_i \to G_0$ are both surjective because of corollary 3.8. Let us give an alternative proof for injectivity. Let $Z_i = \ker(\mathrm{Ad}_i) = \{g \in G_i \text{ st } \mathrm{Ad}(g)X = X, \forall X \in \mathfrak{g}\}$. Since G_i is connected, it is generated by any neighbourhood of the identity in the sense of proposition 2.1; let V_0 be such a neighbourhood. Taking eventually a subset we can suppose that V_0 is a normal coordinate system. So we have

$$g \exp_{G_i}(X) g^{-1} = \exp_{g_i}(X)$$

for every $X \in V_0$. Using proposition 2.1 we deduce that $gxg^{-1} = x$ for every $x \in G_i$, thus $g \in Z(G_i)$. That proves that $\ker(\mathrm{Ad}_i) \subset Z(G_i)$. The assumption of triviality of $Z(G_i)$ concludes injectivity of Ad_i. □

Corollary 3.10.
Let $\mathfrak{g}$ be a real Lie algebra with center $\{0\}$. Then the center of $\mathrm{Int}(\mathfrak{g})$ *is only composed of the identity.*

Proof. We note $G' = \mathrm{Int}(\mathfrak{g})$ and Z his center; ad is the adjoint representation of $\mathfrak{g}$ and Ad', ad', the ones of G' and $\mathrm{ad}(\mathfrak{g})$ respectively. We consider the map $\theta \colon G'/Z \to \mathrm{Int}(\mathrm{ad}(\mathfrak{g}))$, $\theta([g]) = \mathrm{Ad}'(g)$. By the second item of the corollary 3.8, $[g] \to \mathrm{Ad}_{G'}(g)$ is an analytic homomorphism from G' to $\mathrm{Int}(\mathfrak{g}')$ where $\mathfrak{g}'$ is the Lie algebra of G'; this is $\mathrm{ad}(\mathfrak{g})$. So $\theta \colon G'/Z \to \mathrm{Int}(\mathfrak{g}')$ is isomorphic.

Now we consider the map $s \colon \mathfrak{g} \to \mathrm{ad}(\mathfrak{g})$, $s(X) = \mathrm{ad}(X)$; this is an isomorphism. We also consider $S \colon G' \to \mathrm{GL}(\mathrm{ad}(\mathfrak{g}))$, $S(g) = s \circ g \circ s^{-1}$. The Lie algebra of $S(G')$ is $\mathrm{ad}(\mathfrak{g}') = \mathrm{ad}\big(\mathrm{ad}(\mathfrak{g})\big)$. Then $S(G')$ is the subset

[6] Note that $\mathfrak{g}$ and $\mathfrak{h}$ are not groups; by $[X]$, we mean $[X] = \{X + h \text{ st } h \in \mathfrak{h}\}$.

of $\mathrm{GL}(\operatorname{ad}\mathfrak{g})$ whose Lie algebra is $\operatorname{ad}\big(\operatorname{ad}\mathfrak{g}\big)$, i.e. exactly $\mathrm{Int}(\operatorname{ad}\mathfrak{g})$. So S is an isomorphism $S\colon G'\to\mathrm{Int}(\operatorname{ad}\mathfrak{g})$. From all this,

$$S(e^{\operatorname{ad}X})=s\circ e^{\operatorname{ad}X}\circ s^{-1}=e^{\operatorname{ad}'(\operatorname{ad}X)}=\mathrm{Ad}'(e^{\operatorname{ad}X}). \tag{3.19}$$

With this equality, $S^{-1}\circ\theta\colon G'/Z\to G'$ is an isomorphism which sends $[g]$ on g for any $g\in Z$. Then Z can't contains anything else than the identity. ☐

If we relax the assumptions of the trivial center, we have a counter-example with $\mathfrak{g}=\mathbb{R}^3$ and the commutations relation

$$[X_1,X_2]=X_3,\quad [X_1,X_3]=[X_2,X_3]=0.$$

The group $\mathrm{Int}(\mathfrak{g})$ is abelian; then his center is the whole group, although $\mathfrak{g}$ is not abelian.

Note that two groups which have the same Lie algebra are not necessarily isomorphic. For example the sphere S^2 and $\mathbb{R}^2$ both have $\mathbb{R}^2$ as Lie algebra. But two groups with same Lie algebra are locally the same. More precisely, we have the following lemma.

Lemma 3.11.
If G is a Lie group and H, a topological subgroup of G with the same Lie algebra ($\mathfrak{h}=\mathfrak{g}$), then there exists a common neighbourhood A of e of G and G on which the products in G and H are the same.

Proof. The exponential is a diffeomorphism between $U\subset\mathfrak{g}$ and $V\subset G$ and between $U'\subset\mathfrak{h}$ and $W\subset H$ (obvious notations). We consider an open $\mathcal{O}\subset\mathfrak{h}$ such that $\mathcal{O}\subset U\subset U'$. The exponential is diffeomorphic from $\mathcal{O}$ to a certain open A in G and H. Since H is a subgroup of G, the product e^Xe^Y of elements in A is the same for H and G. (cf error **??**) ☐

Under the same assumptions, we can say that H contains at least the whole G_0 because it is generated by any neighbourhood of the identity. Since H is a subgroup, the products keep in H.

For a semisimple Lie group, the Lie algebras $\partial(\mathfrak{g})$ and $\operatorname{ad}(\mathfrak{g})$ are the same. Then $\mathrm{Int}(\mathfrak{g})$ contains at least the identity component of $\mathrm{Aut}(\mathfrak{g})$. Since $\mathrm{Int}(\mathfrak{g})$ is connected, for a semisimple group, it is the identity component of $\mathrm{Aut}(\mathfrak{g})$.

3.3 Killing form

The **Killing form** of $\mathcal{G}$ is the symmetric bilinear form :

$$B(X,Y)=Tr(\operatorname{ad}X\circ\operatorname{ad}Y). \tag{3.20}$$

It is **invariant** in the sense of

$$B\big((\operatorname{ad}S)X,Y\big)=-B\big(X,(\operatorname{ad}S)Y\big), \tag{3.21}$$

$\forall X,\,Y,\,S\in\mathcal{G}$.

Proposition 3.12.
If $\varphi\colon\mathcal{G}\to\mathcal{G}$ is an automorphism of $\mathcal{G}$, then

$$B(\varphi(X),\varphi(Y))=B(X,Y).$$

Proof. The fact that φ is an automorphism of $\mathcal{G}$ is written as $\varphi\circ\operatorname{ad}X=\operatorname{ad}(\varphi(X))\circ\varphi$, or

$$\operatorname{ad}(\varphi(X))=\varphi\circ\operatorname{ad}X\circ\varphi^{-1}.$$

Then

$$\begin{aligned}\mathrm{Tr}(\operatorname{ad}(\varphi(X))\circ\operatorname{ad}(\varphi(Y)))&=\mathrm{Tr}(\varphi\circ\operatorname{ad}X\circ\varphi^{-1}\circ\varphi\operatorname{ad}Y\circ\varphi^{-1})\\&=\mathrm{Tr}(\operatorname{ad}X\circ\operatorname{ad}Y).\end{aligned} \tag{3.22}$$

☐

Remark 3.13.
The Killing 2-form is a map $B\colon\mathcal{G}\times\mathcal{G}\to\mathbb{R}$. When we say that it is preserved by a map $f\colon G\to G$, we mean that it is preserved by df : $B(df\cdot,df\cdot)=B(\cdot,\cdot)$.

An other important property of the Killing form is its bi-invariance.

Theorem 3.14.
The Killing form is bi-invariant on G.

Remark 3.15.
The Killing form is a priori *only defined on $\mathcal{G} = T_eG$. For A, $B \in T_gG$, one naturally defines*

$$B_g(A,B) = B(dL_{g^{-1}}A, dL_{g^{-1}}B). \tag{3.23}$$

This assures the left invariance of B. Now we prove the right invariance.

Proof of theorem 3.14. Because of the left invariance,

$$B(dR_gX, dR_gY) = B(dL_{g^{-1}}dR_gX, dL_{g^{-1}}dR_gY) = B(\operatorname{Ad}_{g^{-1}} X, \operatorname{Ad}_{g^{-1}} Y).$$

But $\operatorname{Ad}_{g^{-1}} = d(\mathbf{Ad}_{g^{-1}})$ and $\mathbf{Ad}_{g^{-1}}$ is an automorphism of G. Thus by lemma 2.7 and proposition 3.12,

$$B\big(\operatorname{Ad}(g^{-1})X, \operatorname{Ad}(g^{-1})Y\big) = B(X,Y). \tag{3.24}$$

□

Lemma 3.16.
Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{i}$ an ideal in $\mathfrak{g}$. Let $B\colon \mathfrak{g}\times\mathfrak{g} \to \mathbb{R}$ be the Killing form on $\mathfrak{g}$ and $B'\colon \mathfrak{i}\times\mathfrak{i} \to \mathbb{R}$, the one of $\mathfrak{i}$. Then $B' = B|_{\mathfrak{i}\times\mathfrak{i}}$, i.e. the Killing form on $\mathfrak{g}$ descent to the ideal $\mathfrak{i}$.

Proof. If W is a subspace of a (finite dimensional) vector space V and $\phi\colon V \to W$ and endomorphism, then $\operatorname{Tr}\phi = \operatorname{Tr}(\phi|_W)$. Indeed, if $\{X_1,\ldots,X_n\}$ is a basis of V such that $\{X_1,\ldots,X_r\}$ is a basis of W, the matrix element ϕ_{kk} is zero for $k > r$. Then

$$\operatorname{Tr}\phi = \sum_{i=1}^{n}\phi_{ii} = \sum_{i=1}^{r}\phi_{ii} = \operatorname{Tr}(\phi|_W).$$

Now consider X, $Y \in \mathfrak{i}$; $(\operatorname{ad}X\circ\operatorname{ad}Y)$ is an endomorphism of $\mathfrak{g}$ which sends $\mathfrak{g}$ to $\mathfrak{i}$ (because $\mathfrak{i}$ is an ideal). Then

$$B'(X,Y) = \operatorname{Tr}\big((\operatorname{ad}X\circ\operatorname{ad}Y)|_{\mathfrak{i}}\big) = \operatorname{Tr}(\operatorname{ad}X\circ\operatorname{ad}Y) = B(X,Y).$$

□

We are not going to (not completely) prove an useful formula for some matrix algebras: $B(X,Y) = 2n\operatorname{Tr}(XY)$ (proposition 3.17). We follow [15]. We consider a simple subalgebra $\mathfrak{g}$ of $\mathfrak{gl}(V)$ for a certain vector space V and a nondegenerate ad-invariant symmetric 2-form f. Then there exists a $S \in \operatorname{GL}(\mathfrak{g})$ such that

$$f(X,Y) = B(SX,Y) \tag{3.25a}$$
$$B(SX,Y) = B(X,SY). \tag{3.25b}$$

If we consider a basis of $\mathfrak{g}$, we can write $f(X,Y)$ (and the Killing) in a matricial form[7] as

$$f(X,Y) = f_{ij}X^iY^j, \qquad B(X,Y) = B_{ij}X^iY^j.$$

Since B is nondegenerate, we can define the matrix (B^{ij}) by $B^{ij}B_{jk} = \delta^i_k$. It is easy to see that the searched endomorphism of $\mathfrak{g}$ is given by $S^k_l = f_{kj}B^{jl}$.

Using the invariance (3.21) of the Killing form and (3.25b), we find

$$B\big((\operatorname{ad}X\circ S)Y, Z\big) = -B\big((S\circ\operatorname{ad}X)Z, Y\big)$$

for any X, Y, $Z \in \mathfrak{g}$. Now using (3.25a),

$$f\big((S^{-1}\circ\operatorname{ad}X\circ S)Y, Z\big) = -f\big((\operatorname{ad}X)Z, Y\big) = f\big((\operatorname{ad}Z)X, Y\big) = f\big(Z, (\operatorname{ad}X)Y\big). \tag{3.26}$$

Since f is nondegenerate, we find $\operatorname{ad}X\circ S = S\circ\operatorname{ad}X$. It follows from Schurs'lemma that $S = \lambda I$. Note that $f(X,Y) = \lambda B(X,Y)$; this proves a certain unicity of the Killing form relatively to his invariance properties.

Now we consider $f(X,Y) = \operatorname{Tr}(XY)$. This is symmetric because of the cyclic invariance of the trace and this is *ad*-invariant because of the formula $\operatorname{Tr}([a,b]c) = \operatorname{Tr}(a[b,c])$ which holds for any matrices a, b, c.

The newt step is to show that f is nondegenerate; we define

$$\mathfrak{g}^\perp = \{X \in \mathfrak{g} \text{ st } f(X,Y) = 0\,\forall Y \in \mathfrak{g}\}.$$

[7] We systematically use the sum convention on the repeated subscript.

The simplicity of $\mathfrak{g}$ ($\mathfrak{g}$ has no proper ideals) makes $\mathfrak{g}$ equal to 0 or $\mathfrak{g}$. Indeed consider $Z \in \mathfrak{g}^\perp$. For any $X, Y \in \mathfrak{g}$, we have

$$0 = f(Z, [X,Y]) = f([Z,X], Y).$$

Then $[Z,X] \in \mathfrak{g}^\perp$ and $\mathfrak{g}^\perp$ is an ideal. We will see that the reality is $\mathfrak{g}^\perp = 0$ (cf. error **??**). Let us suppose $\mathfrak{g}^\perp = \mathfrak{g}$ and consider the lemma 3.52 with $A = B = \mathfrak{g}$. We define

$$M = \{X \in \mathfrak{g} \text{ st } [X, \mathfrak{g}] \subset \mathfrak{g}\} = \mathfrak{g}.$$

If $X \in M$ satisfies $\mathrm{Tr}(XY) = 0$ for any $Y \in M$, then X is nilpotent. Here, $X \in M$ is not a true condition because $M = \mathfrak{g}$. Since $\mathfrak{g}^\perp = \mathfrak{g}$, the trace condition is also trivial. Then $\mathfrak{g}$ is made up with nilpotent endomorphisms of V. Then lemma 3.30 makes all the $X \in \mathfrak{g}$ ad-nilpotent, so that $\mathfrak{g}$ is nilpotent. (cf. remark 3.33)

By the third item of proposition 3.29, $\mathcal{Z}(\mathfrak{g}) \neq 0$ which contradicts the simplicity of $\mathfrak{g}$. Then $\mathfrak{g}^\perp = 0$ and f is nondegenerate. Finally,

$$B(X,Y) = \lambda \,\mathrm{Tr}(X,Y) \tag{3.27}$$

for a certain real number λ. With a certain amount of work (in [8, 15] for example), one can determine the exact value of λ when $\mathfrak{g}$ is the Lie algebra of $n \times n$ matrices with vanishing trace.

Proposition 3.17.
If $\mathfrak{g}$ is the Lie algebra of $n \times n$ matrices with vanishing trace, then

$$B(X,Y) = 2n\,\mathrm{Tr}(XY).$$

3.4 Solvable and nilpotent algebras

If $\mathfrak{g}$ is a Lie algebra, the **derived Lie algebra** is

$$\mathscr{D}\mathfrak{g} = \mathrm{Span}\{[X,Y] \text{ st } X, Y \in \mathfrak{g}\}.$$

We naturally define $\mathscr{D}^0\mathfrak{g} = \mathfrak{g}$ and $\mathscr{D}^n\mathfrak{g} = \mathscr{D}(\mathscr{D}^{n-1}\mathfrak{g})$ this is the **derived series**. Each $\mathscr{D}^n\mathfrak{g}$ is an ideal in $\mathfrak{g}$. We also define the **central decreasing sequence** by $\mathfrak{a}^0 = \mathfrak{a}$, $\mathfrak{a}^{p+1} = [\mathfrak{a}, \mathfrak{a}^p]$.

Definition 3.18.
*The Lie algebra $\mathfrak{g}$ is **solvable** if there exists a $n \geqslant 0$ such that $\mathscr{D}^n\mathfrak{g} = \{0\}$. A Lie group is solvable when its Lie algebra is.*

*The Lie algebra $\mathfrak{g}$ is **nilpotent** if $\mathfrak{g}^n = 0$ for some n. We say that $\mathfrak{g}$ is* ad*-nilpotent if $\mathrm{ad}(X)$ is a nilpotent endomorphism of $\mathfrak{g}$ for each $X \in \mathfrak{g}$.*

Do not confuse *nilpotent* and *solvable* algebras. A nilpotent algebra is always solvable, while the algebra spanned by $\{A, B\}$ with the relation $[A,B] = B$ is solvable but not nilpotent.

If $\mathfrak{g} \neq \{0\}$ is a solvable Lie algebra and if n is the smallest natural such that $\mathscr{D}^n\mathfrak{g} = \{0\}$, then $\mathscr{D}^{n-1}\mathfrak{g}$ is a non zero abelian ideal in $\mathfrak{g}$. We conclude that a solvable Lie algebra is never semisimple (because the center of a semisimple Lie algebra is zero).

A Lie algebra is said to fulfil the **chain condition** if for every ideal $\mathfrak{h} \neq \{0\}$ in $\mathfrak{g}$, there exists an ideal $\mathfrak{h}_1$ in $\mathfrak{h}$ with codimension 1.

Lemma 3.19.
A Lie algebra is solvable if and only if it fulfils the chain condition.

Proof. Necessary condition. The Lie algebra $\mathfrak{g}$ is solvable (then $\mathscr{D}\mathfrak{g} \neq \mathfrak{g}$) and $\mathfrak{h}$ is an ideal in $\mathfrak{g}$. We consider $\mathfrak{h}_1$, a subspace of codimension 1 in $\mathfrak{h}$ which contains $\mathscr{D}\mathfrak{h}$. It is clear that $\mathfrak{h}_1$ is an ideal in $\mathfrak{h}$ because $[H_1, H] \in \mathscr{D}\mathfrak{h} \subset \mathfrak{h}_1$.
Sufficient condition. We have a sequence

$$\{0\} = \mathfrak{g}_n \subset \mathfrak{g}_{n-1} \subset \ldots \subset \mathfrak{g}_0 = \mathfrak{g} \tag{3.28}$$

where $\mathfrak{g}_r$ is an ideal of codimension 1 in $\mathfrak{g}_{r-1}$. Let A be the unique vector in $\mathfrak{g}_{r-1}$ which don't belong to $\mathfrak{g}_r$. When we write $[X,Y]$ with $X, Y \in \mathfrak{g}_{r-1}$, at least one of X or Y is not A (else, it is zero) then at least one of the two is in $\mathfrak{g}_r$. But $\mathfrak{g}_r$ is an ideal; then $[X,Y] \in \mathfrak{g}_r$. Thus $\mathscr{D}(\mathfrak{g}_{r-1}) \subset \mathfrak{g}_r$ and

$$\mathscr{D}^n\mathfrak{g} = \mathscr{D}^{n-1}\mathscr{D}\mathfrak{g} \subset \mathscr{D}^{n-1}\mathfrak{g}_1 \subset \ldots \subset \mathfrak{g}_n = 0.$$

$\square$

Theorem 3.20 (Lie theorem)**.**
Consider $\mathfrak{g}$, a real (resp. complex) solvable Lie algebra and a real (resp. complex) vector space $V \neq \{0\}$. If $\pi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a homomorphism, then there exists a non zero vector in V which is eigenvector of all the elements of $\pi(\mathfrak{g})$.

Problem and misunderstanding 5.
It is strange to be stated for real and complex Lie algebras. Following [8], this is only true for complex Lie algebras while there exists other versions for reals ones.

Proof. Let us do it by induction on the dimension of $\mathfrak{g}$. We begin with $\dim\mathfrak{g} = 1$. In this case, π is just a map $\pi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ such that $\pi(aX) = a\pi(X)$. We have to find an eigenvector for the homomorphism $\pi(X)\colon V \to V$. Such a vector exists from the Jordan decomposition 3.48. Indeed, if there are no eigenvectors, there are no spaces V_i and the decomposition $V = \sum V_i$ can't be true.

Now we consider a general solvable Lie algebra $\mathfrak{g}$ and we suppose that the theorem is true for any solvable Lie algebra with dimension less that $\dim\mathfrak{g}$. Since $\mathfrak{g}$ is solvable, there exists an ideal $\mathfrak{h}$ of codimension 1 in $\mathfrak{g}$; then there exists a $e_0 \neq 0 \in V$ which is eigenvector of all the $\pi(H)$ with $H \in \mathfrak{h}$. So we have $\lambda\colon \mathfrak{h} \to \mathbb{R}$ naturally defined by

$$\pi(H)e_0 = \lambda(H)e_0.$$

Now we consider $X \in \mathfrak{g}\backslash\mathfrak{h}$ and $e_{-1} = 0$, $e_p = \pi(X)^p e_0$ for $p = 1, 2, \ldots$ We will show that $\pi(H)e_p = \lambda(H)e_p$ mod $(e_0, \ldots, e_{p-1})$ for all $H \in \mathfrak{h}$ and $p \geqslant 0$. It is clear for $p = 0$. Let us suppose that it is true for p. Then

$$\begin{aligned}
\pi(H)e_{p+1} &= \pi(H)\pi(X)e_p \\
&= \pi([H, X])e_p + \pi(X)\pi(H)e_p \\
&= \lambda([H, X])e_p + \pi(X)\lambda(H)e_p \\
&\qquad \operatorname{mod}\ (e_0, \ldots, e_{p-1}, \pi(X)e_0, \ldots, \pi(X)e_{p-1}).
\end{aligned} \tag{3.29}$$

But we can put $\pi([H, X])$ and $\pi(X)e_i$ into the modulus. Thus we have

$$\pi(H)e_{p+1} = \lambda(H)e_{p+1} \quad \operatorname{mod}\ (e_0, \ldots, e_p).$$

Now we consider the subspace of V given by $W = \operatorname{Span}\{e_p\}_{p=1,\ldots}$. The algebra $\pi(\mathfrak{h})$ leaves W invariant and our induction hypothesis works on $(\pi(\mathfrak{h}), W)$; then one can find in W a common eigenvector for all the $\pi(H)$. This vector is the one we were looking for. □

Corollary 3.21.
Let $\mathfrak{g}$ be a solvable Lie group and π a representation of $\mathfrak{g}$ on a finite dimensional vector space V. Then there exists a basis $\{e_1, \ldots, e_n\}$ of V in which all the endomorphism $\pi(X)$, $X \in \mathfrak{g}$ are upper triangular matrices.

Proof. Consider $e_1 \neq 0 \in V$, a common eigenvector of all the $\pi(X)$, $X \in \mathfrak{g}$. We consider $E_1 = \operatorname{Span}\{e_1\}$. The representation π induces a representation π_1 of $\mathfrak{g}$ on the space V/E_1. If $V/E_1 \neq \{0\}$, we have a $e_2 \in V$ such that $(e_2 + E_1) \in V/E_1$ is an eigenvector of all the $\pi_(X)$.

In this manner, we build a basis $\{e_1, \ldots, e_n\}$ of V such that $\pi(X)e_i = 0 \mod (e_1, \ldots, e_i)$ for all $X \in \mathfrak{g}$. In this basis, $\pi(X)$ has zeros under the diagonal. □

Theorem 3.22.
Let V be a real or complex vector space and $\mathfrak{g}$, a subalgebra of $\mathfrak{gl}(V)$ made up with nilpotent elements. Then

(i) *$\mathfrak{g}$ is nilpotent;*

(ii) *$\exists v \neq 0$ in V such that $\forall Z \in \mathfrak{g}$, $Zv = 0$;*

(iii) *There exists a basis of V in which the elements of $\mathfrak{g}$ are matrices with only zeros under the diagonal.*

Proof. First item. We consider a $Z \in \mathfrak{g}$ and we have to see that $\operatorname{ad}_{\mathfrak{g}} Z$ is a nilpotent endomorphism of $\mathfrak{g}$. Be careful on a point: an element X of $\mathfrak{g}$ is nilpotent as endomorphism of V while we want to prove that $\operatorname{ad} X$ is nilpotent as endomorphism of $\mathfrak{g}$. We denote by L_Z and R_Z, the left and right multiplication; since we are in a matrix algebra, the bracket is given by the commutator: $\operatorname{ad} Z = L_Z - R_Z$. We have

$$(\operatorname{ad} Z)^p(X) = \sum_{i=0}^{p} (-1)^p \binom{p}{i} Z^{p-i} X Z^i \tag{3.30}$$

There exists a $k \in \mathbb{N}$ such that $Z^k = 0$. For this k, $(\operatorname{ad} Z)^{2k+1}$ is a sum of terms of the form $Z^{p-i}XZ^i$: either $p-i$ either i is always bigger than k. But $\operatorname{ad}_{\mathfrak{g}} Z$ is the restriction of $\operatorname{ad} Z$ (which is defined on $\mathfrak{gl}(V)$) to $\mathfrak{g}$. Then $\mathfrak{g}$ is nilpotent.
Second item. Let $r = \dim \mathfrak{g}$. If $r = 1$, we have only one $Z \in \mathfrak{g}$ and $Z^k = 0$ for a certain (minimal) $k \in \mathbb{N}$. We take v such that $w = Z^{k-1}v \neq 0$ (this exists because k is the minimal natural with $Z^k = 0$). Then $Zw = 0$.

Now we suppose that the claim is valid for any algebra with dimension less than r. Let $\mathfrak{h}$ be a strict subalgebra of $\mathfrak{g}$ with maximal dimension. If $H \in \mathfrak{h}$, $\operatorname{ad}_{\mathfrak{g}} H$ is a nilpotent endomorphism of $\mathfrak{g}$ which sends $\mathfrak{h}$ onto itself. Thus $\operatorname{ad}_{\mathfrak{g}} H$ induces a nilpotent endomorphism H^* on the vector space $\mathfrak{g}/\mathfrak{h}$. We consider the set $\mathcal{A} = \{H^* \text{ st } H \in \mathfrak{h}\}$; this is a subalgebra of $\mathfrak{gl}(\mathfrak{g}/\mathfrak{h})$ made up with nilpotent elements which has dimension strictly less than r.

The induction assumption gives us a non zero $u \in \mathfrak{g}/\mathfrak{h}$ which is sent to 0 by all $\mathcal{A}$, i.e. $(\operatorname{ad}_{\mathfrak{g}} H)u = 0$ in $\mathfrak{g}/\mathfrak{h}$. In other words, $u \in \mathfrak{g}\backslash\mathfrak{h}$ is such that $(\operatorname{ad}_{\mathfrak{g}} H)u \in \mathfrak{h}$.

The space $\mathfrak{h} + \mathbb{K}X$ (here, $\mathbb{K}$ denotes $\mathbb{R}$ or $\mathbb{C}$) of $\mathfrak{g}$ is a subalgebra of $\mathfrak{g}$. Indeed, with obvious notations,

$$[H + kX, H' + k'X] = [H, H'] + \operatorname{ad} H(k'X) - \operatorname{ad} H'(kX) + kk'[X, X]. \tag{3.31}$$

The first term lies in $\mathfrak{h}$ because it is a subalgebra; the second and third therms belongs to $\mathfrak{h}$ by definition of X. The last term is zero. Since $\mathfrak{h}$ is maximal, $\mathfrak{h} + \mathbb{K}X = \mathfrak{g}$. Then (3.31) shows that $\mathfrak{h}$ is also an ideal. Now we consider

$$W = \{e \in V \text{ st } \forall H \in \mathfrak{h}, He = 0\}.$$

Since $\dim \mathfrak{h} < r$, $W \neq \{0\}$ from our induction assumption. Furthermore, for $e \in W$, $HXe = [H, X]e + XHe = 0$. Then $X \cdot W \subset W$. The restriction of X to W is nilpotent. Then there exists a $v \in W$ such that $Xv = 0$. For him $Hv = 0$ because $v \in W$ and $Xv = 0$ by definition of X. Then $Gv = 0$ for any $G \in \mathfrak{h} + \mathbb{K}X = \mathfrak{g}$.
Third item. Let e_1 be a non zero vector in V such that $Ze_1 = 0$ for any $Z \in \mathfrak{g}$ (the existence comes from the second item). We consider $E_1 = \operatorname{Span} e_1$. Any $Z \in \mathfrak{g}$ induces a nilpotent endomorphism Z^* on the vector space V/E_1. If $V/E_1 \neq \{0\}$, we take a $e_2 \in V\backslash E_1$ such that $e_2 + E_1 \in V/E_1$ fulfils $Z^*(e_2 + E_1) = 0$ for all $Z \in \mathfrak{g}$. By going on so, we have $Ze_1 = 0$, $Ze_i = 0 \mod (e_1, \ldots, e_{i-1})$. In this basis, the matrix of Z has zeros on and under the diagonal. □

Corollary 3.23.
Let us consider V, a finite dimensional vector space on $\mathbb{K}$ and $\mathfrak{g}$, a subalgebra of $\mathfrak{gl}(V)$ made up with nilpotent elements. Then if $s \geqslant \dim V$ and $X_i \in \mathfrak{g}$, we have $X_1X_2 \ldots X_s = 0$.

Proof. We write the X_i's in a basis where they have zeros on and under the diagonal. It is rather easy to see that each product push the non zero elements into the upper right corner. □

Corollary 3.24.
A nilpotent algebra is solvable.

Proof. The algebra $\operatorname{ad}_{\mathfrak{g}}(\mathfrak{g})$ is a subalgebra of $\mathfrak{gl}(\mathfrak{g})$ made up with nilpotent endomorphisms of $\mathfrak{g}$. The product of s (see notations of previous corollary) such endomorphism is zero. In particular $\mathfrak{g}$ is solvable. □

We recall the definition of the central decreasing sequence: $\mathfrak{a}^0 = \mathfrak{a}$, $\mathfrak{a}^{p+1} = [\mathfrak{a}, \mathfrak{a}^p]$.

Corollary 3.25.
A Lie algebra $\mathfrak{a}$ is nilpotent if and only if $\mathfrak{a}^m = \{0\}$ for $m \geqslant \dim \mathfrak{a}$.

Proof. The direct sense is easy: we use corollary 3.23 with $\mathfrak{g} = \operatorname{ad}(\mathfrak{a})$ ($\dim \mathfrak{g} = \dim \mathfrak{a}$). Since $\mathfrak{g}$ is nilpotent, for any $X_i \in \mathfrak{g}$ we have $X_1 \ldots X_s$, so that $\mathfrak{a}^m = 0$. The inverse sense is trivial.

□

Corollary 3.26.
A nilpotent Lie algebra $\mathfrak{a} \neq \{0\}$ has a non zero center

Proof. If m is the smallest natural such that $\mathfrak{a}^m = 0$, $\mathfrak{a}^{m-1}$ is in the center. □

Lemma 3.27.
If $\mathfrak{i}$ and $\mathfrak{j}$ are ideals in $\mathfrak{g}$, then we have a canonical isomorphism $\psi\colon (\mathfrak{i} + \mathfrak{j})/\mathfrak{j} \to \mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j})$ given by

$$\psi([x]) = \bar{i}$$

if $x = i + j$ with $i \in \mathfrak{i}$ and $j \in \mathfrak{j}$. Here classes with respect to $\mathfrak{j}$ are denoted by $[.]$ and the one with respect to $(\mathfrak{i} \cap \mathfrak{j})$ by a bar.

Proof. We first have to see that ψ is well defined. If $x' = i + j + j'$, $\psi([x]) = \bar{i}$ because $j + j' \in \mathfrak{j}$. If $x = i' + j'$ (an other decomposition for $x = i + j$), $\bar{i} = \bar{j}$, $j' - j = i - i' \in \mathfrak{j} \cap \mathfrak{i}$. Then $\bar{i} = \overline{i' + j' - j} = \bar{i'}$.

Now it is easy to see that ψ is a homomorphism. □

Proposition 3.28.
Let $\mathfrak{g}$ and $\mathfrak{g}'$ be Lie algebras.

(i) If $\mathfrak{g}$ is solvable then any subalgebra is solvable and if $\phi\colon \mathfrak{g} \to \mathfrak{g}'$ is a Lie algebra homomorphism, then $\phi(\mathfrak{g})$ is solvable in $\mathfrak{g}'$.

(ii) If $\mathfrak{i}$ is a solvable ideal in $\mathfrak{g}$ such that $\mathfrak{g}/\mathfrak{i}$ is solvable, then $\mathfrak{g}$ is solvable.

(iii) If $\mathfrak{i}$ and $\mathfrak{j}$ are solvable ideals in $\mathfrak{g}$, then $\mathfrak{i} + \mathfrak{j}$ is also a solvable ideal in $\mathfrak{g}$.

Proof. First item. If $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, then $\mathscr{D}^k\mathfrak{h} \subset \mathscr{D}^k\mathfrak{g}$, so that $\mathfrak{h}$ is solvable. Now consider $\mathfrak{h} = \phi(\mathfrak{g}) \subset \mathfrak{g}'$. This is a subalgebra of $\mathfrak{g}'$ because $[h, h'] = [\phi(g), \phi(g')] = \phi([g, g']) \in \mathfrak{h}$. It is clear that $\mathscr{D}(\phi(\mathfrak{g})) \subset \phi(\mathscr{D}(\mathfrak{g}))$ and

$$\mathscr{D}^2(\phi(\mathfrak{g})) = \mathscr{D}\big(\mathscr{D}\phi(\mathfrak{g})\big) \subset \mathscr{D}(\phi\mathscr{D}(\mathfrak{g})) \subset \phi\mathscr{D}\mathscr{D}(\mathfrak{g}) = \phi(\mathscr{D}^2(\mathfrak{g})). \tag{3.32}$$

Repeating this argument, $\mathscr{D}^k(\mathfrak{h}) \subset \phi(\mathscr{D}^k\mathfrak{g})$. So $\mathfrak{h}$ is also solvable. Note that $\phi([g, g']) = [\phi(g), \phi(g')] \subset \mathscr{D}(\pi(\mathfrak{g}))$. Then

$$\mathscr{D}^k\pi(\mathfrak{g}) = \pi(\mathscr{D}^k\mathfrak{g}). \tag{3.33}$$

Second item. Let n be the smallest integer such that $\mathscr{D}^n(\mathfrak{g}/\mathfrak{i}) = 0$; we look at the canonical homomorphism $\pi\colon \mathfrak{g} \to \mathfrak{g}/\mathfrak{i}$. This satisfies $\mathscr{D}^n(\pi(\mathfrak{g})) = \pi(\mathscr{D}^n\mathfrak{g}) = 0$. Then $\mathscr{D}^n(\mathfrak{g}) \subset \mathfrak{i}$. If $\mathscr{D}^m\mathfrak{i} = 0$, then $\mathscr{D}^{m+n}\mathfrak{g} = 0$.
Third item. The space $\mathfrak{i}/(\mathfrak{i} \cap \mathfrak{j})$ is the image of $\mathfrak{i}$ by a homomorphism, then it is solvable and $(\mathfrak{i} + \mathfrak{j})/\mathfrak{j}$ is also solvable. The second item makes $\mathfrak{i} + \mathfrak{j}$ solvable. □

Now we consider $\mathfrak{g}$, any Lie algebra and $\mathfrak{s}$ a maximum solvable ideal i.e. it is included in none other solvable ideal. Let us consider $\mathfrak{i}$, an other solvable ideal in $\mathfrak{g}$. Then $\mathfrak{i} + \mathfrak{s}$ is a solvable ideal; since $\mathfrak{s}$ is maximal, $\mathfrak{i} + \mathfrak{s} = \mathfrak{s}$. Thus there exists an unique maximal solvable ideal which we call the **radical** of $\mathfrak{g}$. It will be often denoted by $\operatorname{Rad}\mathfrak{g}$. If β is a symmetric bilinear form, his **radical** is the set

$$S = \{x \in \mathfrak{g} \text{ st } \beta(x, y) = 0\ \forall y \in \mathfrak{g}\}. \tag{3.34}$$

The form β is nondegenerate if and only if $S = \{0\}$.

Proposition 3.29.
Let $\mathfrak{g}$ and $\mathfrak{g}'$ be Lie algebras.

(i) If $\mathfrak{g}$ is nilpotent, then his subalgebras are nilpotent and if $\phi\colon \mathfrak{g} \to \mathfrak{g}'$ is a Lie algebra homomorphism, then $\phi(\mathfrak{g})$ is nilpotent.

(ii) If $\mathfrak{g}/\mathcal{Z}(\mathfrak{g})$ is nilpotent, then $\mathfrak{g}$ is nilpotent. For recall,

$$\mathcal{Z}(\mathfrak{g}) = \{z \in \mathfrak{g} \textit{ st } [x, z] = 0\ \forall x \in \mathfrak{g}\}.$$

(iii) If $\mathfrak{g}$ is nilpotent, then $\mathcal{Z}(\mathfrak{g}) \neq 0$.

Proof. The proof of the first item is the same as the one of 3.28. Now if $(\mathfrak{g}/\mathcal{Z}(\mathfrak{g}))^n = 0$, then $\mathfrak{g}^n/\mathcal{Z}(\mathfrak{g}) = 0$; thus $\mathfrak{g}^n \subset \mathcal{Z}(\mathfrak{g})$, so that $\mathfrak{g}^{n+1} = [\mathfrak{g}, \mathcal{Z}(\mathfrak{g})] = 0$. Finally, if n is the smallest natural such that $\mathfrak{g}^n = 0$, then $[\mathfrak{g}^{n-1}, \mathfrak{g}] = 0$ and $\mathfrak{g}^{n-1} \subset \mathcal{Z}(\mathfrak{g})$. □

The condition to be nilpotent can be reformulated by $\exists n \in \mathbb{N}$ such that $\forall X_i$, $Y \in \mathfrak{g}$,

$$(\operatorname{ad} X_1 \circ \ldots \circ \operatorname{ad} X_n)Y = 0,$$

in particular for any $X \in \mathfrak{g}$, there exists a $n \in \mathbb{N}$ such that $(\operatorname{ad} X)^n = 0$. An element for which such a n exists is **ad-nilpotent**. If $\mathfrak{g}$ is nilpotent, then all his elements are ad-nilpotent.

Some results without proof :

Lemma 3.30.
If $X \in \mathfrak{gl}(V)$ is a nilpotent endomorphism, then $\operatorname{ad} X$ is nilpotent.

Remark 3.31.
The inverse implication is not true, as the unit matrix shows.

Theorem 3.32 (Engel)**.**
A Lie algebra is nilpotent if and only if all his elements are ad-nilpotent.

For a proof see [8].

Remark 3.33.
The combination of these two last results makes that if $\mathfrak{g} \subset \mathfrak{gl}(V)$ *is made up with nilpotent endomorphisms of* V*, then* $\mathfrak{g}$ *is nilpotent as Lie algebra.*

3.5 Flags and nilpotent Lie algebras

Here we give a "flag description" of some previous results. In particular the chain (3.28). If V is a vector space of dimension $n < \infty$, a **flag** in V is a chain of subspaces $0 = V_0 \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n = \mathfrak{g}$ with $\dim V_k = k$. If $x \in \operatorname{End} V$ fulfils $x(V_i) \subset V_i$, then we say that x **stabilise** the flag.

Theorem 3.34.
If $\mathfrak{g}$ *is a subalgebra of* $\mathfrak{gl}(V)$ *in which the elements are nilpotent endomorphisms and if* $V \neq 0$*, then there exists a* $v \in V$*,* $v \neq 0$ *such that* $\mathfrak{g}v = 0$*.*

Proof. This is the second item of theorem 3.22. □

Corollary 3.35.
Under the same assumptions, there exists a flag (V_i) *stable under* $\mathfrak{g}$ *such that* $\mathfrak{g}V_i \subset V_{i-1}$*. In other words, there exists a basis of* V *in which the matrices of* $\mathfrak{g}$ *are nilpotent; this basis is the one given by the flag.*

Proof. Let $v \neq 0$ such that $\mathfrak{g}v = 0$ which exists by the theorem and $V_1 = \operatorname{Span} v$. We consider $W = V/V_1$; the action of $\mathfrak{g}$ on W is also made up with nilpotent endomorphisms. Then we go on with V_1 and $W_1 = W/V_2$,... □

Lemma 3.36.
If $\mathfrak{g}$ *is nilpotent and if* $\mathfrak{i}$ *is an non trivial ideal in* $\mathfrak{g}$*, then* $\mathfrak{i} \cap \mathcal{Z}(\mathfrak{g}) \neq 0$*.*

Proof. Since $\mathfrak{i}$ is an ideal, $\mathfrak{g}$ acts on $\mathfrak{i}$ with the adjoint representation. The restriction of an element $\operatorname{ad} X$ for $X \in \mathfrak{g}$ to $\mathfrak{i}$ is in fact a nilpotent element in $\mathfrak{gl}(\mathfrak{i})$. Then we have a $I \in \mathfrak{i}$ such that $\mathfrak{g}I = 0$. Thus $I \in \mathfrak{i} \cap \mathcal{Z}(\mathfrak{g})$. □

Theorem 3.37.
Let $\mathfrak{g}$ *be a solvable Lie subalgebra of* $\mathfrak{gl}(V)$*. If* $V \neq 0$*, then* V *posses a common eigenvector for all the endomorphisms of* $\mathfrak{g}$*.*

Proof. This is exactly the Lie theorem 3.20 □

Corollary 3.38 (Lie theorem)**.**
Let $\mathfrak{g}$ *be a solvable subalgebra of* $\mathfrak{gl}(V)$*. Then* $\mathfrak{g}$ *stabilize a flag of* V*.*

Proof. This corollary is the corollary given in 3.21.

We consider v_1 the vector given by theorem 3.37. Since it is eigenvector of all $\mathfrak{g}$, $\operatorname{Span} v_1$ is stabilised by $\mathfrak{g}$. Next we consider v_2 in the complementary which is also a common eigenvector,... □

Corollary 3.39.
If $\mathfrak{g}$ *is a solvable Lie algebra, then there exists a chain of ideals in* $\mathfrak{g}$

$$0 = \mathfrak{g}_0 \subset \mathfrak{g}_1 \subset \ldots \subset \mathfrak{g}_n = \mathfrak{g}$$

with $\dim \mathfrak{g}_k = k$*.*

Proof. If $\phi\colon \mathfrak{g} \to \mathfrak{gl}(V)$ is a finite-dimensional representation of $\mathfrak{g}$, then $\phi(\mathfrak{g})$ is solvable by proposition 3.29. Then $\phi(\mathfrak{g})$ stabilises a flag of V. Now we take as ϕ the adjoint representation of $\mathfrak{g}$. A stable flag is the chain of ideals; indeed if $\mathfrak{g}_i$ is a part of the flag, then $\forall H \in \mathfrak{g}$ $\operatorname{ad} H \mathfrak{g}_i \subset \mathfrak{g}_i$ because the flag is invariant. □

Corollary 3.40.
If $\mathfrak{g}$ *is solvable then* $X \in \mathscr{D}\mathfrak{g}$ *implies that* $\operatorname{ad}_{\mathfrak{g}} X$ *is nilpotent. In particular* $\mathscr{D}\mathfrak{g}$ *is nilpotent.*

Proof. We consider the ideals chain of previous corollary and an adapted basis: $\{X_1,\ldots,X_n\}$ is such that $\{X_1,\ldots,X_i\}$ spans $\mathfrak{g}_i$. In such a basis the matrices of $\mathrm{ad}(\mathfrak{g})$ are upper triangular and it is easy to see that in this case, the matrices of $[\mathrm{ad}\,\mathfrak{g},\mathrm{ad}\,\mathfrak{g}]$ are *strictly* upper triangular: they have zeros on the diagonal. But $[\mathrm{ad}\,\mathfrak{g},\mathrm{ad}\,\mathfrak{g}]=\mathrm{ad}_{\mathfrak{g}}[\mathfrak{g},\mathfrak{g}]$. Then for $X\in\mathrm{ad}_{\mathfrak{g}}\mathscr{D}\mathfrak{g}$, $\mathrm{ad}_{\mathfrak{g}}X$ is nilpotent. *A fortiori*, $\mathrm{ad}_{\mathscr{D}\mathfrak{g}}X$ is nilpotent and by the theorem of Engel 3.32, $\mathscr{D}\mathfrak{g}$ is nilpotent. □

The following lemma is computationally useful because it says that if X is a nilpotent element of a Lie algebra, then $g\cdot X$ is also nilpotent with (at most) the same order.

Lemma 3.41.
The following formula
$$\mathrm{ad}(g\cdot X)^nY=g\cdot\mathrm{ad}(X)^n(g^{-1}\cdot Y)\tag{3.35}$$
holds for all $g\in G$ and $X,Y\in\mathfrak{g}$,

The proof is a simple induction on n.

3.6 Semisimple Lie algebras

A useful reference to go trough semisimple Lie algebras is [16]. Very few proofs, but the statements of all the useful results with explanations.

Definition 3.42.
A Lie algebra is ***semisimple*** *if it has no proper abelian invariant Lie subalgebra. A Lie algebra is* ***simple*** *if it is not abelian and has no proper Lie subalgebra.*

In that definition, we say that a Lie subalgebra $\mathfrak{h}$ is **invariant** if $\mathrm{ad}(\mathfrak{g})\mathfrak{h}\subset\mathfrak{h}$.

There are a lot of equivalent characterisations. Here are some that are going to be proved (or not) in the next few pages. A Lie algebra is semisimple if an only if one of the following conditions is respected.

(i) The Killing form is nondegenerate.

(ii) The radical of $\mathfrak{g}$ is zero (theorem 3.56).

(iii) There are no abelian proper invariant subalgebra.

Problem and misunderstanding 6.
I think that in the following I took the degenerateness of Killing as definition.

The Killing form is a convenient way to define a Riemannian metric on a semisimple[8] Lie group.

Corollary 3.43.
An automorphism of a semisimple Lie group is an isometry for the Killing metric. Stated in other words,
$$\mathrm{Aut}(G)\subset \mathtt{Iso}\,G.\tag{3.36}$$

Proof. By lemma 2.7, if f is an automorphism of G, df is an automorphism of $\mathcal{G}$. Now, by proposition 3.12, f is an isometry of G. □

Proposition 3.44.
Let $\mathfrak{g}$ be a semisimple Lie algebra, $\mathfrak{a}$ an ideal in $\mathfrak{g}$, and $\mathfrak{a}^\perp=\{X\in\mathfrak{g}\ st\ B(X,A)=0\forall A\in\mathfrak{a}\}$. Then

(i) $\mathfrak{a}^\perp$ *is an ideal,*

(ii) $\mathfrak{g}=\mathfrak{a}\oplus\mathfrak{a}^\perp$,

(iii) $\mathfrak{a}$ *is semisimple,*

Proof. First item. We have to show that for any $X\in\mathfrak{g}$ and $P\in\mathfrak{a}^\perp$, $[X,P]\in\mathfrak{a}^\perp$, or $\forall\,Y\in\mathfrak{a}$, $B(Y,[X,P])=0$. From invariance of B,
$$B(Y,[X,P])=B(P,[Y,X])=0.$$
Second item. Since B is nondegenerate, $\dim\mathfrak{a}+\dim\mathfrak{a}^\perp=\dim\mathfrak{g}$. Let us consider $Z\in\mathfrak{g}$ and X, $Y\in\mathfrak{a}\cap\mathfrak{a}^\perp$. We have $B(Z,[X,Y])=B([Z,X],Y)=0$. Then $[X,Y]=0$ because $B(Z,[X,Y])=0$ for any Z and B is nondegenerate. Thus $\mathfrak{a}\cap\mathfrak{a}^\perp$ is abelian. It is also an ideal because $\mathfrak{a}$ and $\mathfrak{a}^\perp$ are.

[8] In this case, B is nondegenerate.

Now we consider $\mathfrak{b}$, a complementary of $\mathfrak{a}\cap\mathfrak{a}^\perp$ in $\mathfrak{g}$, $Z \in \mathfrak{g}$ and $T \in \mathfrak{a}\cap\mathfrak{a}^\perp$. The endomorphism $E = \operatorname{ad} T\circ \operatorname{ad} Z$ sends $\mathfrak{a}\cap\mathfrak{a}^\perp$ to $\{0\}$. Indeed consider $A \in \mathfrak{a}\cap\mathfrak{a}^\perp$; $(\operatorname{ad} Z)A \in \mathfrak{a}\cap\mathfrak{a}^\perp$ because it is an ideal, and then $(\operatorname{ad} T\circ \operatorname{ad} Z)A = 0$ because it is abelian.

The endomorphism E also sends $\mathfrak{b}$ to $\mathfrak{a}\cap\mathfrak{a}^\perp$ (it may not be surjective); then $\operatorname{Tr}(\operatorname{ad} T\circ \operatorname{ad} Z) = 0$ and $\mathfrak{a}\cap\mathfrak{a}^\perp = \{0\}$. Since B is nondegenerate, $\dim\mathfrak{a} + \dim\mathfrak{a}^\perp = \dim\mathfrak{g}$. Then $\mathfrak{a}\oplus\mathfrak{a}^\perp = \mathfrak{g}$ is well a direct sum.
Third item. From lemma 3.16, the Killing form of $\mathfrak{g}$ descent to the ideal $\mathfrak{a}$; then it is also nondegenerate and $\mathfrak{a}$ is also semisimple. □

Corollary 3.45.
A semisimple Lie algebra has center $\{0\}$.

Proof. If $Z \in \ker\mathfrak{g}$, $\operatorname{ad} Z = 0$. So $B(Z, X) = 0$ for any $X \in \mathfrak{g}$. Since B is nondegenerate, it implies $Z = 0$. □

Corollary 3.46.
If $\mathfrak{g}$ is a semisimple Lie algebra, it can be written as a direct sum

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_r$$

where the $\mathfrak{g}_i$ are simples ideals in $\mathfrak{g}$. Moreover each simple ideal in $\mathfrak{g}$ is a direct sum of some of them.

Proof. If $\mathfrak{g}$ is simple, the statement is trivial. If it is not, we consider $\mathfrak{a}$, an ideal in $\mathfrak{g}$. Proposition 3.44 makes $\mathfrak{g} = \mathfrak{a}\oplus\mathfrak{a}^\perp$. Since $\mathfrak{a}$ and $\mathfrak{a}^\perp$ are semisimple, we can once again brake them in the same way. We do it until we are left with simple algebras.

For the second part, consider $\mathfrak{b}$ a simple ideal in $\mathfrak{g}$ which is not a sum of $\mathfrak{g}_i$. Then $[\mathfrak{g}_i, \mathfrak{b}] \subset \mathfrak{g}_i \cap \mathfrak{b} = \{0\}$. Then $\mathfrak{b}$ is in the center of $\mathfrak{g}$. This contradict corollary 3.45. □

Proposition 3.47.
If $\mathfrak{g}$ is semisimple then

$$\operatorname{ad}(\mathfrak{g}) = \partial(\mathfrak{g}),$$

i.e. any derivation is an inner automorphism :

Proof. We saw at page 40 that $\operatorname{ad}(\mathfrak{g}) \subset \partial(\mathfrak{g})$ holds without assumptions of (semi)simplicity. Now we consider D, a derivation: $\forall X \in \mathfrak{g}$,

$$\operatorname{ad}(DX) = [D, \operatorname{ad} X].$$

Then $\operatorname{ad}(\mathfrak{g})$ is an ideal in $\partial(\mathfrak{g})$ because the commutator of $\operatorname{ad} X$ with any element of $\partial(\mathfrak{g})$ still belongs to $\operatorname{ad}(\mathfrak{g})$. Let us denote by $\mathfrak{a}$ the orthogonal complement of $\operatorname{ad}(\mathfrak{g})$ in $\partial(\mathfrak{g})$ (for the Killing metric). The algebra $\operatorname{ad}(\mathfrak{g})$ is semisimple because of it isomorphic to $\mathfrak{g}$. Since the Killing form on $\operatorname{ad}(\mathfrak{g})$ is nondegenerate, $\mathfrak{a} \cap \operatorname{ad}(\mathfrak{g}) = \{0\}$. Finally $D \in \mathfrak{a}$ implies $[D, \operatorname{ad} X] \in \mathfrak{a} \cap \operatorname{ad}(\mathfrak{g}) = \{0\}$. Then $\operatorname{ad}(DX) = 0$ for any $X \in \mathfrak{g}$, so that $D = 0$. This shows that $\mathfrak{a} = \{0\}$, so that $\operatorname{ad}(\mathfrak{g}) = \partial(\mathfrak{g})$. □

If V is a finite dimensional space, a subspace W in V is **invariant** under a subset $G \subset \operatorname{Hom}(V, V)$ if $sW \subset W$ for any $s \in G$. The space V is **irreducible** when V and $\{0\}$ are the only two invariant subspaces. The set G is **semisimple** if any invariant subspace has an invariant complement. In this case, the vector space split into $V = \sum_i V_i$ with V_i invariant and irreducible.

Theorem 3.48 (Jordan decomposition)**.**
Any element $A \in \operatorname{Hom}(V, V)$ is decomposable in one and only one way as $A = S + N$ with S semisimple and N nilpotent and $NS = SN$. Furthermore, S and N are polynomials in A. More precisely :

If V is a complex vector space and $A \in \operatorname{Hom}(V, V)$ with $\lambda_1, \ldots, \lambda_r$ his eigenvalues, we pose

$$V_i = \{v \in V \ st\ (A - \lambda_i\mathbb{1})^k v = 0 \text{ for large enough } k\}.$$

Then

(i) $V = \sum_{i=1}^r V_i$,

(ii) *each V_i is invariant under A,*

(iii) *the semisimple part of A is given by*

$$S(\sum_{i=1}^r v_i) = \sum_{i=1}^r \lambda_i v_i,$$

for $v_i \in V_i$,

(iv) the characteristic polynomial of A is

$$\det(\lambda\mathbb{1} - A) = (\lambda - \lambda_1)^{d_1} \ldots (\lambda - \lambda_r)^{d_r}$$

where $d_i = \dim V_i$ $(1 \leqslant i \leqslant r)$.

3.6.1 Jordan decomposition

If V is a finite dimensional vector space, we say that an element of $\operatorname{End} V$ is **semisimple** when it is diagonalisable. We know that two commuting semisimple endomorphism are simultaneously diagonalisable. So the sums and differences of semisimple elements still are semisimple.

Let E_{kl} be the $(n+2) \times (n+2)$ matrix with a 1 at position (k, l) and 0 anywhere else: $(E_{kl})_{ij} = \delta_{ki}\delta_{lj}$. An easy computation show that

$$E_{kl}E_{ab} = \delta_{la}E_{kb}, \tag{3.37}$$

and

$$[E_{kl}, E_{rs}] = \delta_{lr}E_{ks} - \delta_{sk}E_{rl}. \tag{3.38}$$

Now we give a great theorem without proof.

Theorem 3.49 (Jordan decomposition)**.**
Let V be a finite dimensional vector space and $x \in \operatorname{End} V$.

(i) There exists one and only one choice of $x_s, x_n \in \operatorname{End}(V)$ such that $x = x_s + x_n$, x_s is semisimple, n_n is nilpotent and $[x_s, x_n] = 0$.

(ii) There exists polynomials p and q without independent term such that $x_s = p(x)$, $x_n = q(x)$; in particular if $y \in \operatorname{End} V$ commutes with x, then it commutes with x_s and x_n.

(iii) If $A \subset B \subset V$ are subspaces of V and if $x(B) \subset A$, then $x_s(B) \subset A$ and $x_n(B) \subset A$.

As an example consider the adjoint representation of $\mathfrak{gl}(V)$. As seen in lemma 3.30, if $x \in \mathfrak{gl}(V)$ is nilpotent, then $\operatorname{ad} x$ is also nilpotent.

Lemma 3.50.
If $x \in \mathfrak{gl}(V)$ is semisimple, then $\operatorname{ad} x$ is also semisimple.

Proof. We choose a basis $\{v_1, \cdots, v_n\}$ of V in which x is diagonal with eigenvalues $a_1, \ldots, a_n$. For $\mathfrak{gl}(V)$, we consider the basis $\{E_{ij}\}$ in which E_{ij} is the matrix with a 1 at position (i, j) and zero anywhere else. This satisfies $[E_{kl}, E_{rs}] = \delta_{lr}E_{ks} - \delta_{sk}E_{rl}$. We easily check that $E_{kl}(v_i) = \delta_{li}v_k$. Since we are in a matrix algebra, the adjoint action is the commutator: $(\operatorname{ad} x)E_{ij} = [x, E_{ij}]$; as we know that $x = a_k E_{kk}$,

$$(\operatorname{ad} x)E_{ij} = a_k[E_{kk}, E_{ij}] = (a_i - a_j)E_{ij} \tag{3.39}$$

which proves that $\operatorname{ad} x$ has a diagonal matrix in the basis $\{E_{ij}\}$ of $\mathfrak{gl}(V)$. Furthermore, we have an explicit expression for his matrix: the eigenvalues are $(a_i - a_j)$.

□

Lemma 3.51.
Let $x \in \operatorname{End} V$ with his Jordan decomposition $x = x_s + x_n$. Then the Jordan decomposition of $\operatorname{ad} x$ is

$$\operatorname{ad} x = \operatorname{ad} x_s + \operatorname{ad} x_n. \tag{3.40}$$

Proof. We already know that $\operatorname{ad} x_s$ is semisimple and $\operatorname{ad} x_n$ is nilpotent. They commute because $[\operatorname{ad} x_s, \operatorname{ad} x_n] = \operatorname{ad}[x_s, x_n] = 0$. Then the unicity part of Jordan theorem 3.49 makes (3.40) the Jordan decomposition of $\operatorname{ad} x$. □

3.6.2 Cartan criterion

Let us recall a result: $\mathscr{D}\mathfrak{g} = \mathfrak{g}^1$, $[\mathscr{D}\mathfrak{g}, \mathscr{D}\mathfrak{g}] \subset \mathfrak{g}^2$; then $\mathscr{D}^k\mathfrak{g} \subset \mathfrak{g}^k$. Thus if $\mathfrak{g}$ is nilpotent, it is solvable. On the other hand, by the Engel theorem 3.32, $\mathscr{D}\mathfrak{g}$ is nilpotent if and only if all the $\operatorname{ad}_{\mathscr{D}\mathfrak{g}} x$ are nilpotent for $x \in \mathscr{D}\mathfrak{g}$.

Lemma 3.52.
Let $A \subset B$ be two subspace of $\mathfrak{gl}(V)$ with $\dim V < \infty$. We pose
$$M = \{x \in \mathfrak{gl}(V) \text{ st } [x, B] \subset A\},$$
and we suppose that $x \in M$ verify $\operatorname{Tr}(x \circ y) = 0$ for all $y \in M$. Then x is nilpotent.

Proof. We use the Jordan decomposition $x = x_s + x_n$ and a basis in which x_s takes the form $diag(a_1, \ldots, a_m)$; let $\{v_1, \ldots, v_m\}$ be this basis. We denotes by E the vector space on $\mathbb{Q}$ spanned by $\{a_1, \ldots, a_m\}$. We want to prove that $x_s = 0$, i.e. $E = 0$. Since E has finite dimension, it is equivalent to prove that its dual is zero. In other words, we have to see that any linear map $f : E \to \mathbb{Q}$ is zero.

We consider $y \in \mathfrak{gl}(V)$, an element whose matrix is $diag(f(a_1), \ldots, f(a_m))$ and (E_{ij}), the usual basis of $\mathfrak{gl}(V)$. We know that
$$(\operatorname{ad} x_s)E_{ij} = (a_i - a_j)E_{ij}, \tag{3.41a}$$
$$(\operatorname{ad} y)E_{ij} = (f(a_i) - f(a_j))E_{ij}. \tag{3.41b}$$
It is always possible to find a polynomial r on $\mathbb{R}$ without constant term such that $r(a_i - a_j) = f(a_i) - f(a_j)$. Note that this is well defined because of the linearity of f : if $a_i - a_j = a_k - a_l$, then $f(a_i) - f(a_j) = f(a_k) - f(a_l)$. Since $\operatorname{ad} x_s$ is diagonal, $r(\operatorname{ad} x_s)$ is the matrix with $r(\operatorname{ad} x_s)_{ii}$ on the diagonal and zero anywhere else. Then $r(\operatorname{ad} x_s) = \operatorname{ad} y$. By lemma 3.51, $\operatorname{ad} x_s$ is the semisimple part of $\operatorname{ad} x$, then $\operatorname{ad} y$ is a polynomial without constant term with respect to $\operatorname{ad} x$ (second point of theorem 3.49).

Since $(\operatorname{ad} y)B \subset A$, $y \in M$ and $\operatorname{Tr}(xy) = 0$. It is easy to convince ourself that the s_n part of x will not contribute to the trace because x_n is strictly upper triangular and y is diagonal. From the explicit forms of x_s and y,
$$\operatorname{Tr}(xy) = \sum_i a_i f(a_i) = 0.$$
This is a $\mathbb{Q}$-linear combination of element of E : we have to see it as a_i being a basis vector and $f(a_i)$ a coefficient, so that we can apply f on both sides to find $0 = \sum_i f(a_i)^2$. Then for all i, $f(a_i) = 0$, so that $f = 0$ because the a_i spans E. □

Theorem 3.53 (Cartan criterion)**.**
Let $\mathfrak{g}$ be a subalgebra of $\mathfrak{gl}(V)$. We suppose that $\operatorname{Tr}(xy) = 0$ $\forall x \in \mathscr{D}\mathfrak{g}, y \in \mathfrak{g}$. Then $\mathfrak{g}$ is solvable.

Proof. It is sufficient to prove that $\mathscr{D}\mathfrak{g}$ is nilpotent indeed if we write $\mathscr{D}^k\mathfrak{g} \subset \mathfrak{g}^k$ with $\mathscr{D}\mathfrak{g}$ instead of $\mathfrak{g}$, $\mathscr{D}^{k+1}\mathfrak{g} \subset (\mathscr{D}\mathfrak{g})^k$. If $\mathscr{D}\mathfrak{g}$ is nilpotent, $(\mathscr{D}\mathfrak{g})^n = 0$ and $\mathscr{D}^{n+1}\mathfrak{g} = 0$ so that $\mathfrak{g}$ is solvable.

Let us consider $x \in \mathscr{D}\mathfrak{g}$. We have to prove that it is ad-nilpotent (see the Engel theorem 3.32). Let $A = \mathscr{D}\mathfrak{g}$, $B = \mathfrak{g}$ and $M = \{x \in \mathfrak{gl}(V) \text{ st } [x\mathfrak{g}] \subset \mathscr{D}\mathfrak{g}\}$. By definition of $\mathscr{D}\mathfrak{g}$, $\mathfrak{g} \subset M$. The lemma 3.52 will conclude that $x \in \mathscr{D}\mathfrak{g}$ is nilpotent if $\operatorname{Tr}(xy) = 0$ for any $y \in M$. Here we just have this equality for $y \in \mathfrak{g}$.

A typical generator of $\mathscr{D}\mathfrak{g}$ is $[x, y]$ with x, $y \in \mathfrak{g}$. Take a $z \in M$; by the formula $\operatorname{Tr}([x, y]z) = \operatorname{Tr}(x[y, z])$, the trace that we have to check is
$$\operatorname{Tr}([x, y]z) = \operatorname{Tr}(x[y, z]) = \operatorname{Tr}([y, z]x). \tag{3.42}$$
But with $z \in M$, $[y, z] \in \mathscr{D}\mathfrak{g}$, then $\operatorname{Tr}([x, y]z) = \operatorname{Tr}([y, z]x) = 0$. Thus we are in the situation of the lemma. □

Corollary 3.54.
A Lie algebra $\mathfrak{g}$ for which $\operatorname{Tr}(\operatorname{ad} x \circ \operatorname{ad} y) = 0$ for all $x \in \mathscr{D}\mathfrak{g}$, $y \in \mathfrak{g}$ is solvable.

Proof. We consider $\mathfrak{h} = \operatorname{ad}\mathfrak{g}$; this is a subalgebra of $\mathfrak{gl}(V)$ such that $a \in \mathscr{D}\mathfrak{h}$ and $b \in \mathfrak{h}$ imply $\operatorname{Tr}(ab) = 0$. In order to see it, remark that $a \in \mathscr{D}\mathfrak{h}$ can be written as $a = [\operatorname{ad} x, \operatorname{ad} y] = \operatorname{ad}[x, y]$ for certain x, $y \in \mathfrak{g}$. Then $\operatorname{Tr}(ab) = \operatorname{Tr}(\operatorname{ad}[x, y] \operatorname{ad} z)$ with x, y, $z \in \mathfrak{g}$; this is zero from the hypothesis. Then $\mathfrak{h} = \operatorname{ad}\mathfrak{g}$ is solvable.

It is also known that $\ker(\operatorname{ad}) = \mathcal{Z}(\mathfrak{g})$ is also solvable. Now we consider $\mathfrak{m}$ a complementary of $\mathcal{Z}(\mathfrak{g})$ in $\mathfrak{g}$: $\mathfrak{g} = \mathcal{Z} \oplus \mathfrak{m}$. The Lie algebra $\operatorname{ad}(\mathfrak{m})$ is solvable and the homomorphism ϕ: $\operatorname{ad}\mathfrak{m} \to \mathfrak{m}$ defined by $\phi(\operatorname{ad} x) = x$ is well defined. From the first item of the proposition 3.28, $\mathfrak{m}$ is solvable. With obvious notations, an element of $\mathscr{D}\mathfrak{m}$ can be written as $[m, m']$ (because $\mathcal{Z}(\mathfrak{g})$ don't contribute to $\mathscr{D}\mathfrak{g}$). Then $\mathscr{D}\mathfrak{g} = \mathscr{D}\mathfrak{m}$, so that $\mathfrak{g}$ is as much solvable than $\mathfrak{m}$. □

Lemma 3.55.
The radical of a Lie algebra is non zero if and only if it has at least non zero abelian ideal.

Proof. The radical of $\mathfrak{g}$ is its unique maximal solvable ideal. An eventually non empty abelian ideal should be in the radical.

Let us now consider that the radical is non zero, and consider the derived series of $\operatorname{Rad}\mathfrak{g}$. Since $\operatorname{Rad}\mathfrak{g}$ is solvable, we can consider n, the minimal integer such that $\mathscr{D}^n \operatorname{Rad}\mathfrak{g} = 0$. Then $\mathscr{D}^{n-1} \operatorname{Rad}\mathfrak{g}$ is a non zero abelian ideal. □

Theorem 3.56.
A Lie algebra is semisimple if and only if its radical is zero.

Proof. Direct sense. We suppose $\operatorname{Rad}\mathfrak{g} = 0$ and we consider S, the radical of the Killing form :

$$S = \{X \in \mathfrak{g} \text{ st } B(X,Y) = 0\,\forall Y \in \mathfrak{g}\}.$$

By definition, for any $X \in S$ and $Y \in \mathfrak{g}$, $\operatorname{Tr}(\operatorname{ad}X \circ \operatorname{ad}Y) = 0$. The Cartan criterion makes $\operatorname{ad}S$ solvable and the corollary 3.54 makes S solvable.

Now, the *ad*-invariance of the Killing form turns S into an ideal, so that $S \subset \operatorname{Rad}(\mathfrak{g})$ because any solvable ideal is contained in $\operatorname{Rad}\mathfrak{g}$. From the assumptions, $\operatorname{Rad}S = 0$, then $S \subset \operatorname{Rad}\mathfrak{g} = 0$. This shows that the Killing form is nondegenerate.

Inverse sense. We suppose $S = 0$ and we will show that any abelian ideal of $\mathfrak{g}$ is in S. In this case, if A is a solvable ideal with $\mathscr{D}^n A = 0$, then $\mathscr{D}^{n-1}A$ is an abelian ideal, so that $\mathscr{D}^{n-1}A = 0$. By induction, $A = 0$.

Let I be an abelian ideal of $\mathfrak{g}$, $X \in I$ and $Y \in \mathfrak{g}$. Then $\operatorname{ad}X \circ \operatorname{ad}Y$ is nilpotent because for $Z \in \mathfrak{g}$,

$$(\operatorname{ad}X \operatorname{ad}Y \operatorname{ad}X \operatorname{ad}Y)Z = (\operatorname{ad}X \operatorname{ad}Y)\underbrace{([X,[Y,Z]])}_{=X_1 \in I} = (\operatorname{ad}X)\underbrace{[Y,X_1]}_{=X_2 \in I} = (\operatorname{ad}X)X_2 = 0. \tag{3.43}$$

Then $0 = \operatorname{Tr}(\operatorname{ad}X \operatorname{ad}Y) = B(X,Y)$ and $X \in S$, so that $I \subset S = 0$. □

3.6.3 More about radical

If $\mathfrak{g}$ is a Lie algebra whose radical is $\mathfrak{r}$, we say that a subalgebra $\mathfrak{s}$ of $\mathfrak{g}$ is a **Levi subalgebra** if $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{s}$.

Any Lie algebra posses a Levi subalgebra[9].

Lemma 3.57.
If $\mathfrak{a}$ is an ideal in a Lie algebra $\mathfrak{g}$, then

$$\operatorname{Rad}\mathfrak{a} = (\operatorname{Rad}\mathfrak{r}) \cap \mathfrak{a}.$$

Before to begin the proof, let us recall that lemma 3.27 gives us an isomorphism $\psi\colon (\mathfrak{a}+\mathfrak{b})/\mathfrak{a} \to \mathfrak{b}/(\mathfrak{a}\cap\mathfrak{b})$ when $\mathfrak{a}$ and $\mathfrak{b}$ are ideals in $\mathfrak{g}$.

Proof of the lemma. If $\mathfrak{r}$ is the radical of $\mathfrak{g}$, then the radical of $\mathfrak{g}/\mathfrak{r}$ is zero, so that $\mathfrak{r}/\mathfrak{r}$ is semisimple. Let $\mathfrak{a}$ be an ideal in $\mathfrak{g}$, then $(\mathfrak{a}+\mathfrak{r})/\mathfrak{r}$ is an ideal in the semisimple Lie algebra $\mathfrak{g}\mathfrak{r}$, so that it is also semisimple. From the isomorphism, $\mathfrak{a}/(\mathfrak{a}\cap\mathfrak{r})$ is also semisimple and $\mathfrak{a}\cap\mathfrak{r}$ must contains the radical of $\mathfrak{a}$. Indeed if a solvable ideal of $\mathfrak{a}$ where not in $\mathfrak{a}\cap\mathfrak{r}$, then this should give rise to a non zero solvable ideal in $\mathfrak{a}/(\mathfrak{a}\cap\mathfrak{r})$ although the latter is semisimple. Then $\mathfrak{a}\cap\mathfrak{r} = \operatorname{Rad}\mathfrak{a}$. □

Proposition 3.58.
If A is a compact group of automorphisms of the Lie algebra $\mathfrak{g}$, then there exists a Levi subalgebra of $\mathfrak{g}$ which is invariant under A.

Proof. Let $\mathfrak{r}$ be the radical of $\mathfrak{g}$; we will split our proof into two cases following $[\mathfrak{r},\mathfrak{r}] = 0$ or not.

The radical is abelian. In this first case we consider an induction with respect to the dimension of $\mathfrak{g}$. We consider $\bar{\mathfrak{g}} = \mathfrak{g}/[\mathfrak{r},\mathfrak{r}]$ and $\bar{\mathfrak{r}} = \mathfrak{r}/[\mathfrak{r},\mathfrak{r}]$: these are algebras with one less dimension that $\mathfrak{g}$ and $\mathfrak{r}$. We denote by $\pi\colon \mathfrak{g} \to \bar{\mathfrak{g}}$ the natural projection.

We begin to prove that $\bar{\mathfrak{r}}$ is the radical of $\bar{\mathfrak{g}}$. It is clear from the Lie algebra structure on a quotient that $\bar{\mathfrak{r}}$ is an ideal because $\mathfrak{r}$ is. It is also clear that $\bar{\mathfrak{r}}$ is solvable. We just have to see that $\bar{\mathfrak{r}}$ is maximal in $\bar{\mathfrak{g}}$. For this, suppose that $\bar{\mathfrak{r}} \cup \overline{X}$ is a solvable ideal in $\bar{\mathfrak{g}}$. Then it is easy to see that $\mathfrak{r} \cup X$ is an ideal in $\mathfrak{g}$. Taking commutators in $\bar{\mathfrak{r}} \cup \overline{X}$, we always finish in $\overline{0} \in \bar{\mathfrak{g}}$, i.e. in $[\mathfrak{r},\mathfrak{r}]$. Taking again some commutators, we finish on $0 \in \mathfrak{g}$ because $\mathfrak{r}$ is solvable. This contradict the maximality of $\mathfrak{r}$.

Since A is made up of automorphisms, it leaves $\mathfrak{r}$ invariant, so that it also acts on $\bar{\mathfrak{g}}$ as an automorphism group: $a\overline{X} = \overline{aX}$ for $a \in A$ and $X \in \mathfrak{g}$. From the induction assumption, we can find a Levi subalgebra $\bar{\mathfrak{s}}$ in $\bar{\mathfrak{g}}$: $\bar{\mathfrak{s}} \oplus \bar{\mathfrak{r}} = \bar{\mathfrak{g}}$. In this case, the radical of $\pi^{-1}(\bar{\mathfrak{s}})$ is $[\mathfrak{r},\mathfrak{r}]$. Indeed in the one hand, $\bar{\mathfrak{r}} \cap \bar{\mathfrak{s}} = 0$, so that $\pi^{-1}(\bar{\mathfrak{r}} \cap \bar{\mathfrak{s}}) = [\mathfrak{r},\mathfrak{r}]$. In the other hand $\pi^{-1}(\bar{\mathfrak{r}} \cap \bar{\mathfrak{s}}) = \pi^{-1}(\bar{\mathfrak{r}}) \cap \pi^{-1}(\bar{\mathfrak{s}}) = \mathfrak{r} \cap \pi^{-1}(\bar{\mathfrak{s}})$. The lemma 3.57 conclude that $\operatorname{Rad}\pi^{-1}(\bar{\mathfrak{s}}) = [\mathfrak{r},\mathfrak{r}]$.

Now A is a compact group of automorphism which leaves invariant $\pi^{-1}(\bar{\mathfrak{s}})$, so we have a Levi subalgebra $\mathfrak{s}$ of $\pi^{-1}(\bar{\mathfrak{s}})$ invariant under A. We will see that this is in fact a Levi subalgebra of the whole $\mathfrak{g}$, i.e. we have to prove that $\mathfrak{s} \oplus \mathfrak{r} = \mathfrak{g}$. From the definition of $\mathfrak{s}$,

$$\mathfrak{s} \oplus [\mathfrak{r},\mathfrak{r}] = \pi^{-1}(\bar{\mathfrak{s}}),$$

[9] Reference needed.

and by definition of $\bar{\mathfrak{s}}$,

$$\bar{\mathfrak{s}} \oplus \frac{\mathfrak{r}}{[\mathfrak{r},\mathfrak{r}]} = \bar{\mathfrak{g}}.$$

Then

$$\mathfrak{g} = \pi^{-1}(\bar{\mathfrak{s}}) \oplus \mathfrak{r} + [\mathfrak{r},\mathfrak{r}] = \mathfrak{s} \oplus [\mathfrak{r},\mathfrak{r}] \oplus \mathfrak{r} + [\mathfrak{r},\mathfrak{r}] = \mathfrak{s} \oplus \mathfrak{r}. \tag{3.44}$$

We can now pass to the second case: $[\mathfrak{r},\mathfrak{r}] = 0$.

The radical is not abelian. Let $\mathfrak{s}_0$ and $\mathfrak{s}$ be Levi subalgebras of $\mathfrak{g}$. For $X \in \mathfrak{s}_0$, we write

$$X = f(X) + X_{\mathfrak{s}}$$

with respect to the decomposition $\mathfrak{g} = \mathfrak{s} \oplus \mathfrak{r}$. This defines a linear map $f\colon \mathfrak{s}_0 \to \mathfrak{r}$. For any $X, Y \in \mathfrak{s}_0$, $[X_{\mathfrak{s}}, X_{\mathfrak{s}}] = [X,Y] - [X,f(y)] - [f(X),Y]$ because $\mathfrak{r}$ is abelian. Since[10], $[X_{\mathfrak{s}}, X_{\mathfrak{s}}] = [X,Y]_{\mathfrak{s}}$,

$$f([X,Y]) = [X,f(Y)] - [f(X),Y]. \tag{3.45}$$

Now let us consider a map $f\colon \mathfrak{s}_0 \to \mathfrak{r}$ which satisfy this equation. Then the map $X \to X - f(X)$ is an isomorphism between $\mathfrak{s}_0$ and his image which is a Levi subalgebra of $\mathfrak{g}$. Indeed

$$\begin{aligned}
[X,Y] &\to [X,Y] - f([X,Y]) \\
&= [X,Y] - [X,f(Y)] - [f(X),Y] \\
&= [X - f(X), Y - f(Y)].
\end{aligned} \tag{3.46}$$

Now we consider V, the space of all the linear maps $\mathfrak{s}_0 \to \mathfrak{r}$ which fulfil the condition (3.45). We have a bijection between V and the Levi subalgebras of $\mathfrak{g}$: for any Levi subalgebra we associate the map $f \in V$ given by $X = f(X) + X_{\mathfrak{s}}$.

So our proof can be reduced to find a fixed point of V under the action of A. In order to do that, we will see that A is a group of *affine* transformations on V. Consider a $\alpha \in A$ and f_0, f_0^α, f^α be the elements of V corresponding to $\mathfrak{s}_0$, $\mathfrak{s}$ and $\alpha(\mathfrak{s})$. We take a $X \in \mathfrak{s}_0$ and we denote by $\overline{\alpha}(X)$ the $\mathfrak{s}_0$-component of $\alpha(X)$ with respect to the decomposition $\mathfrak{g} = \mathfrak{r} \oplus \mathfrak{s}_0$:

$$\alpha(X) = \overline{\alpha}(X) + \beta(X).$$

This also defines $\beta\colon \mathfrak{g} \to \mathfrak{r}$ and $-\beta(X)$ is the $\mathfrak{r}$-component of $\overline{\alpha}(X)$ with respect to $\mathfrak{g} = \mathfrak{r} \oplus \alpha(\mathfrak{s}_0)$. Since f_0^α just correspond to this decomposition, $f_0^\alpha(\overline{\alpha}(X)) = -\beta(X)$, so that

$$\begin{aligned}
\overline{\alpha}(X) &= f_0^\alpha(\overline{\alpha}(X)) + \alpha(X) \\
&= f_0^\alpha(\overline{\alpha}(X)) + \alpha(f(X)) - \alpha(f(X)) + \alpha(X).
\end{aligned} \tag{3.47}$$

Since $X - f(X) \in \mathfrak{s}$, $\alpha(X) - \alpha(f(X)) \in \alpha(\mathfrak{s})$, then $f_0^\alpha(\overline{\alpha}(X)) + \alpha(X)$ is the $\mathfrak{r}$-component of $\overline{\alpha}(X)$ with respect to $\mathfrak{g} = \mathfrak{r} \oplus \alpha(\mathfrak{s})$. Then

$$f_0^\alpha(\overline{\alpha}(X)) + \alpha(f(X)) = f^\alpha(\overline{\alpha}(X)) = f^\alpha(\overline{\alpha}(X)).$$

Since X was taken arbitrary, $f^\alpha = f_0^\alpha + \alpha \circ f \circ \overline{\alpha}^{-1}$. Then the map $V \to V$, $f \to f^\alpha$ is an affine transformation with translation equals to f_0^α and linear part being $f \to \alpha \circ f \circ \overline{\alpha}$.

A general result shows that a compact group of affine transformations on a vector space has a fixed point. □

3.6.4 Compact Lie algebra

We consider $\mathfrak{g}$, a real Lie algebra and $\mathfrak{h}$, a subalgebra of $\mathfrak{g}$. Let K^* be the analytic subgroup of $\mathrm{Int}(\mathfrak{g})$ which corresponds to the subalgebra $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{h})$ of $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g})$.

Definition 3.59.
*We say that $\mathfrak{h}$ is **compactly embedded** in $\mathfrak{g}$ if K^* is compact. A Lie algebra is **compact** when it is compactly embedded in itself.*

The analytic subgroup of $\mathrm{Int}(\mathfrak{g})$ which corresponds to $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{g})$, by definition, is $\mathrm{Int}(\mathfrak{g})$. Then the compactness of $\mathfrak{g}$ is the one of $\mathrm{Int}(\mathfrak{g})$.

Remark 3.60.
The compactness notion on a Lie group is defined from the topological structure of the Lie group seen as a manifold. It is all but trivial that the compactness on a Lie group is related to the compactness on its Lie algebra; the proposition 3.65 will however make the two notions related in the natural way.

[10] C'est pas clair pourquoi on a ạ.

Remark 3.61.
The topology on K^* is not necessary the same as the induced one from $\mathrm{Int}(\mathfrak{g})$ *and* $\mathrm{Int}(\mathfrak{g})$ *has also not necessary the induced topology from* $\mathrm{GL}(\mathfrak{g})$. *However the next proposition will show that the compactness notion is well the one induced from* $\mathrm{GL}(\mathfrak{g})$.

Proposition 3.62.
We consider $\widetilde{K}$, the same set and group as K^*, but with the induced topology from $\mathrm{GL}(\mathfrak{g})$. *Then $\widetilde{K}$ is compact if and only if K^* is compact.*

Note however that K^* and $\widetilde{K}$ are not automatically the same as manifold.

Proof. K^* compact implies $\widetilde{K}$ compact. The identity map $\iota\colon K^*\to \mathrm{GL}(\mathfrak{g})$ is analytic, and then is continuous because $\mathrm{Int}(\mathfrak{g})$ is by definition an analytic subgroup of $\mathrm{GL}(\mathfrak{g})$ and K^* an analytic subgroup of $\mathrm{Int}(\mathfrak{g})$. If we have a covering of $\widetilde{K}$ with open set $\mathcal{O}_i\cap\widetilde{K}$ of $\widetilde{K}$ ($\mathcal{O}_i$ is open in $\mathrm{GL}(\mathfrak{g})$), the continuity of ι make the finite subcovering of K^* good for $\widetilde{K}$.
$\widetilde{K}$ compact implies K^* compact. If $\widetilde{K}$ is compact, then it is closed in $\mathrm{GL}(\mathfrak{g})$. As set, K^* is closed in $\mathrm{GL}(\mathfrak{g})$ and by definition it is connected. Then by the theorem 2.36, K^* is a topological subgroup of $\mathrm{GL}(\mathfrak{g})$. Consequently, K^* and $\widetilde{K}$ are homeomorphic and they have same topology. □

A lemma without proof[11].

Lemma 3.63.
If G is a compact group in $\mathrm{GL}(n,\mathbb{R})$, *then there exists a G-invariant quadratic form on $\mathbb{R}^n$.*

Proposition 3.64.
Let $\mathfrak{g}$ be a real Lie algebra.

(i) If $\mathfrak{g}$ is semisimple, then $\mathfrak{g}$ is compact if and only if the Killing form is strictly negative definite.

(ii) If it is compact then it is a direct sum
$$\mathfrak{g}=\mathcal{Z}\oplus[\mathfrak{g},\mathfrak{g}] \tag{3.48}$$
where $\mathcal{Z}$ is the center of $\mathfrak{g}$ and the ideal $[\mathfrak{g},\mathfrak{g}]$ is compact and semisimple.

Proof. If the Killing form is nondegenerate. We consider $\mathfrak{g}$, a Lie algebra whose Killing form is strictly negative definite. Up to some dilatations (and a sign), this is the euclidian metric. Then $O(B)$, the group of linear transformations which leave B unchanged is compact in the topology of $\mathrm{GL}(\mathfrak{g})$: this is almost the rotations. From equation (3.36), $\mathrm{Aut}(\mathfrak{g})\subset O(B)$. With this, $\mathrm{Aut}(\mathfrak{g})$ is closed in a compact, then it is compact. Then $\mathrm{Int}(\mathfrak{g})$ is closed in $\mathrm{Aut}(\mathfrak{g})$ –here is the assumption of semi-simplicity– and $\mathrm{Int}(\mathfrak{g})$ is compact.
If $\mathfrak{g}$ is compact. Since $\mathfrak{g}$ is compact, $\mathrm{Int}(\mathfrak{g})$ is compact in the topology of $\mathrm{Aut}(\mathfrak{g})$; then there exists an $\mathrm{Int}(\mathfrak{g})$-invariant quadratic form Q. In a suitable basis $\{X_1,\ldots,X_n\}$ of $\mathfrak{g}$, we can write this form as
$$Q(X)=\sum x_i^2$$
for $X=\sum x_iX_i$. In this basis the elements of $\mathrm{Int}(\mathfrak{g})$ are orthogonal matrices and the matrices of $\mathrm{ad}(\mathfrak{g})$ are skew-symmetric matrices (the Lie algebra of orthogonal matrices). Let us consider a $X\in\mathfrak{g}$ and denote by $a_{ij}(X)$ the matrix of $\mathrm{ad}(X)$. We have
$$B(X,X)=\mathrm{Tr}(\mathrm{ad}\,X\circ\mathrm{ad}\,X)=\sum_i\sum_j a_{ij}(X)a_{ji}(X)=-\sum_{ij}a_{ij}(X)^2\leqslant 0. \tag{3.49}$$

Then the Killing form is negative definite[12]. On the other hand, $B(X,X)=0$ implies $\mathrm{ad}(X)=0$ and $X\in\mathcal{Z}(\mathfrak{g})$. Thus $\mathfrak{g}^{\perp}\subset\mathcal{Z}$. If $\mathfrak{g}$ is semisimple, this center is zero; this conclude the first item of the proposition.
Now $\mathcal{Z}$ is an ideal and corollary 3.46 decomposes $\mathfrak{g}$ as
$$\mathfrak{g}=\mathcal{Z}\oplus\mathfrak{g}'. \tag{3.50}$$

Let us suppose that the restriction of B to $\mathfrak{g}'\times\mathfrak{g}'$ is actually the Killing form on $\mathfrak{g}'$ (we will prove it below). Then the Killing form on $\mathfrak{g}'$ is strictly negative definite; then $\mathfrak{g}'$ is compact.

[11] J'ai même pas trouvé d'énoncé de ce théorème.
[12] Here we use "negative definite" and "*strictly* negative definite"; in some literature, the terminology is slightly different and one says "*semi* negative definite" and "negative definite".

Now we prove that the Killing form on $\mathfrak{g}$ descent to the Killing form on $\mathfrak{g}'$. Remark that $\mathcal{Z}$ is invariant under all the automorphism. Indeed consider $Z \in \mathcal{Z}$, i.e. $[X, Z] = 0$. If σ is an automorphism,

$$[X, \sigma Z] = \sigma[\sigma^{-1}X, Z] = 0.$$

Here the difference between $\mathrm{Int}(\mathfrak{g})$ and $\mathrm{Aut}(\mathfrak{g})$ is the fact that $\mathrm{Int}(\mathfrak{g})$ is compact; then we can construct a $\mathrm{Int}(\mathfrak{g})$-invariant quadratic form Q, but not a $\mathrm{Aut}(\mathfrak{g})$-invariant one. We consider an orthogonal complement (with respect to Q) $\mathfrak{g}'$ of $\mathcal{Z}$:

$$\mathfrak{g} = \mathfrak{g}' \oplus_{\perp} \mathcal{Z}. \tag{3.51}$$

The algebra $\mathfrak{g}'$ is also invariant because for any $Z \in \mathcal{Z}$,

$$Q(Z, \sigma X) = Q(\sigma^{-1}(Z), X) = 0.$$

It is also clear that $\mathcal{Z}$ is invariant under $\operatorname{ad}\mathfrak{g}$ because $(\operatorname{ad}X)Z = 0$. Finally $\mathfrak{g}'$ is invariant as well under $\operatorname{ad}(\mathfrak{g})$. Indeed $a \in \operatorname{ad}(\mathfrak{g})$ can be written as $a = a'(0)$ for a path $a(t) \in \mathrm{Int}(\mathfrak{g})$. We identify $\mathfrak{g}$ and his tangent space (as vector spaces),

$$aX = \frac{d}{dt}\Big[a(t)X\Big]_{t=0}.$$

If $X \in \mathfrak{g}'$, $a(t)X \in \mathfrak{g}'$ for any t because $\mathfrak{g}'$ is invariant under $\mathrm{Int}(\mathfrak{g})$[13]. Thus $a(t)X$ is a path in $\mathfrak{g}'$ and his derivative is a vector in $\mathfrak{g}'$.

All this make $\mathfrak{g}'$ an ideal in $\mathfrak{g}$; then the Killing form descent by lemma 3.16. Now if $X \in \mathfrak{g}$, we have

$$B(X, X) = \mathrm{Tr}(\operatorname{ad}X \circ \operatorname{ad}X) = \sum_{ij} a_{ij}(X)a_{ji}(X) = -\sum_{ij} a_{ij}(X)^2; \tag{3.52}$$

then $B(X, X) \leqslant 0$ and the equality holds if and only if $\operatorname{ad}X = 0$ i.e. if and only if $X \in \mathcal{Z}$. Thus B is strictly negative definite on $\mathfrak{g}'$.

Up to now we have proved that $\mathfrak{g}'$ is semisimple (because B is nondegenerate) and compact (because B is strictly negative definite).

It remains to be proved that $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}] = \mathscr{D}(\mathfrak{g})$. From corollary 3.46, $\mathscr{D}\mathfrak{g}$ has a complementary $\mathfrak{a}$ which is also an ideal: $\mathfrak{g} = \mathscr{D}\mathfrak{g} + \mathfrak{a}$. Then $[\mathfrak{g}, \mathfrak{a}] \subset \mathscr{D}\mathfrak{g}$ and $[\mathfrak{g}, \mathfrak{a}] \subset \mathfrak{a} \cap \mathscr{D}\mathfrak{g} : \{0\}$. Then $\mathfrak{a} \subset \mathcal{Z}$, so that

$$\mathfrak{g} = \mathcal{Z} + \mathscr{D}\mathfrak{g} \qquad \text{(non direct sum).} \tag{3.53}$$

Now we have to prove that the sum is actually direct. The ideal $\mathcal{Z}$ has a complementary ideal $\mathfrak{b} : \mathfrak{g} = \mathcal{Z} \oplus \mathfrak{b}$ and

$$\mathscr{D}\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \subset \underbrace{[\mathfrak{g}, \mathcal{Z}]}_{=0} + [\mathfrak{g}, \mathfrak{b}] \subset \mathfrak{b}.$$

Then $\mathscr{D}\mathfrak{g} \subset \mathfrak{b}$ which implies that $\mathscr{D}\mathfrak{g} \cap \mathcal{Z} = \{0\}$ because the sum $\mathfrak{g} = \mathcal{Z} \oplus \mathfrak{b}$ is direct. Then the sum (3.53) is direct.

□

Proposition 3.65.
A real Lie algebra $\mathfrak{g}$ is compact if and only if one can find a compact Lie group G which Lie algebra is isomorphic to $\mathfrak{g}$.

Proof. Direct sense. Since $\mathfrak{g}$ is compact, $\mathfrak{g} = \mathcal{Z} \oplus \mathscr{D}\mathfrak{g}$ with $\mathscr{D}\mathfrak{g} = \mathfrak{g}'$ compact and semisimple; in particular, the center of $\mathfrak{g}'$ is $\{0\}$. Since $\mathcal{Z}$ is compact and abelian, it is isomorphic to the torus $S^1 \times \ldots \times S^1$. Since $\mathfrak{g}'$ is compact, $\mathrm{Int}(\mathfrak{g}')$ is compact, but the Lie algebra if $\mathrm{Int}(\mathfrak{g}')$ is –by definition– $\operatorname{ad}(\mathfrak{g}')$. The center of a semisimple Lie algebra is zero; then $\operatorname{ad}X' = 0$ implies $X = 0$ (for $X \in \mathfrak{g}'$). Then ad is an isomorphism between $\mathfrak{g}'$ and $\operatorname{ad}\mathfrak{g}'$.

All this shows that –up to isomorphism– $\mathcal{Z}$ and $[\mathfrak{g}, \mathfrak{g}]$ are Lie algebras of compact groups. We know from lemma 1.11 that the Lie algebra of $G \times H$ is $\mathfrak{g} \oplus \mathfrak{h}$. Thus, here, $\mathfrak{g}$ is the Lie algebra of the compact group $S^1 \times \ldots \times S^1 \times \mathrm{Int}(\mathfrak{g})$.
Reverse sense. We consider a compact group G and we have to see the its Lie algebra $\mathfrak{g}$ is compact. If G is connected, Ad_G is an analytic homomorphism from G to $\mathrm{Int}(\mathfrak{g})$. If G is not connected, the Lie algebra of G is T_eG_0 (G_0 is the identity component of G) where G_0 is connected and compact because closed in a compact. □

Proposition 3.66.
Let $\mathfrak{g}$ be a real Lie algebra and $\mathcal{Z}$, the center of $\mathfrak{g}$. We consider $\mathfrak{k}$, a compactly embedded in $\mathfrak{g}$. If $\mathfrak{k} \cap \mathcal{Z} = \{0\}$ then the Killing form of $\mathfrak{g}$ is strictly negative definite on $\mathfrak{k}$.

[13] As physical interpretation, if something is invariant under a group of transformations, it is invariant under the infinitesimal transformations as well.

Proof. Let B be the Killing form on $\mathfrak{g}$ and K the analytic subgroup of $\mathrm{Int}(\mathfrak{g})$ whose Lie algebra is $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{k})$. By assumption, K is a compact Lie subgroup of $\mathrm{GL}(\mathfrak{g})$. Then there exists a quadratic form on $\mathfrak{g}$ invariant under K, and a basis in which the endomorphisms $\mathrm{ad}_{\mathfrak{g}}(T)$ for $T \in \mathfrak{k}$ are skew-symmetric because the matrices of K are orthogonal. If the matrix of $\mathrm{ad}\, T$ is (a_{ij}), then

$$B(T,T) = \sum_{ij} a_{ij}(T) a_{ji}(T) = -\sum_{ij} a_{ij}^2(T) \leqslant 0, \tag{3.54}$$

and the equality hold only if $\mathrm{ad}\, T = 0$ i.e. if $T \in \mathcal{Z}$. From the assumptions, $\mathfrak{k} \cap \mathcal{Z} = \{0\}$; then $B(T,T) = 0$ if and only if $T = 0$. □

3.7 Cartan subalgebras in complex Lie algebras

About Cartan algebra, one can read [8, 11, 12, 17].

In this section $\mathfrak{g}$ will always denotes a complex finite dimensional Lie algebra.

Definition 3.67.
*When $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, the **centralizer** of $\mathfrak{h}$ is the set*

$$\mathcal{Z}(\mathfrak{h}) = \{x \in \mathfrak{g} \text{ st } [x, \mathfrak{h}] \subset \mathfrak{h}\}. \tag{3.55}$$

More generally if $\mathfrak{g}$ is a Lie algebra and if $\mathfrak{a}$, $\mathfrak{b}$ are two subset of $\mathfrak{g}$, the centraliser of $\mathfrak{a}$ in $\mathfrak{b}$ is

$$\mathcal{Z}_{\mathfrak{b}}(\mathfrak{a}) = \{X \in \mathfrak{b} \text{ st } [X, \mathfrak{a}] = 0\}. \tag{3.56}$$

*If $\mathfrak{a}$ is a subalgebra of $\mathfrak{g}$, its **normalizer** is*

$$\mathfrak{n}_{\mathfrak{a}} = \{X \in \mathfrak{g} \text{ st } [X, \mathfrak{a}] \subset \mathfrak{a}\}. \tag{3.57}$$

One can check that $\mathfrak{a}$ is an ideal in $\mathfrak{n}_{\mathfrak{a}}$.

Definition 3.68.
*A subalgebra $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is a **Cartan subalgebra** if it is nilpotent and if it is its own centralizer: $[x, \mathfrak{h}] \subset \mathfrak{h}$ implies $x \in \mathfrak{h}$.*

Our first task is to show that every Lie algebra has a Cartan algebra.

Lemma 3.69 (Primary decomposition theorem)**.**
Let V be a complex vector space and $A\colon V \to V$ be linear map. Then we have the direct sum decomposition

$$V = \bigoplus_{\lambda \in \mathbb{C}} V_\lambda(A) \tag{3.58}$$

where $V_\lambda(A) = \{v \text{ st } (A - \lambda \mathbb{1})^n v = 0 \text{ for some } n \in \mathbb{N}\}$

This is the result that restricts ourself to *complex* Lie algebras when proving that Cartan subalgebras exist. Notice that the sum in (3.58) is reduced to the eigenvalues of A since $\mathfrak{g}_\lambda(A) = 0$ when λ is not an eigenvalue. Indeed if $\big(A - \lambda\mathbb{1}\big)^n Y = 0$ then $(A - \lambda\mathbb{1})^{n-1} Y$ is an eigenvector for A with eigenvalue λ.

For any $\lambda \in \mathbb{C}$ and $X \in \mathfrak{g}$ we consider the space

$$\mathfrak{g}_\lambda(X) = \{Y \in \mathfrak{g} \text{ st } \big(\mathrm{ad}(X) - \lambda\mathbb{1}\big)^n Y = 0 \text{ for some } n\}. \tag{3.59}$$

The primary decomposition theorem implies the decomposition

$$\mathfrak{g} = \bigoplus_\lambda \mathfrak{g}_\lambda(X) \tag{3.60}$$

for each $X \in \mathfrak{g}$.

Lemma 3.70.
For each $X \in \mathfrak{g}$ and $\lambda, \mu \in \mathbb{C}$ we have

$$\big[\mathfrak{g}_\lambda(X), \mathfrak{g}_\mu(X)\big] \subset \mathfrak{g}_{\lambda+\mu}(X). \tag{3.61}$$

Proof. Let $X_\lambda \in \mathfrak{g}_\lambda(X)$ and $X_\mu \in \mathfrak{g}_\mu(X)$. Using the fact that $\mathrm{ad}(X)$ is a derivation we have

$$\mathrm{ad}(X)[X_\lambda, X_\mu] - (\lambda+\mu)[X_\lambda, X_\mu] = \Big[\big(\mathrm{ad}(X)-\mu\mathbb{1}\big)X_\lambda, X_\mu\Big] + \Big[X_\lambda, \big(\mathrm{ad}(X)+\mu\mathbb{1}\big)X_\mu\Big] \tag{3.62}$$

and by induction[14] we find

$$\big(\mathrm{ad}\,Z - (\lambda+\mu)\mathbb{1}\big)^n[X_\lambda, X_\mu] = \sum_{i=0}^{\infty}\binom{n}{i}[(\mathrm{ad}\,Z-\lambda I)^i X_\lambda, (\mathrm{ad}\,Z-\mu I)^{n-i}X\mu] \tag{3.63}$$

which vanishes when n is large enough. □

We say that X is **regular** if $\dim\mathfrak{g}_0(X)$ is the smallest with respect to the others $\dim\mathfrak{g}_0(Y)$.
The following proposition shows that every complex Lie algebra has a Cartan Lie subalgebra.

Proposition 3.71.
If X is regular in $\mathfrak{g}$ then the subalgebra $\mathfrak{g}_0(X)$ is Cartan.

Proof. Since $X \in \mathfrak{g}_0(X)$ we have $\mathrm{ad}(X)\mathfrak{g}_\lambda(X) \subset \mathfrak{g}_\lambda(X)$. Thus we see $\mathrm{ad}(X)$ as a linear operator on $\mathfrak{g}_\lambda(X)$. The operator $\mathrm{ad}(X)|_{\mathfrak{g}_\lambda(X)}$ is nonsingular[15] when $\lambda \neq 0$. Indeed all the eigenvalues of $\mathrm{ad}(X)$ on $\mathfrak{g}_\lambda(X)$ are equal to λ because

$$\big(\mathrm{ad}(X)-\mu\mathbb{1}\big)Y = 0 \tag{3.64}$$

implies $Y \in \mathfrak{g}_\mu(X)$. If $Y \in \mathfrak{g}_\lambda(X)$ it only occurs when $\mu = \lambda$ since the sum (3.58) is direct.

For each eigenvalue λ we have a neighborhood $\mathcal{U}_\lambda$ of X in $\mathfrak{g}_0(X)$ such that for all $Y \in \mathcal{U}_\lambda$, $\mathrm{ad}(Y)$ is nonsingular on $\mathfrak{g}_\lambda(X)$. We consider $\mathcal{U} = \bigcap_\lambda \mathcal{U}_\lambda$ which is a non empty open set since the intersection is taken over the eigenvalues of $\mathrm{ad}(X)$ that are in finite numbers.

Let us prove that the restriction to $\mathfrak{g}_0(X)$ of the linear operator $\mathrm{ad}(Y)$ is nilpotent for each $Y \in \mathcal{U}$. First we have

$$\mathfrak{g}_0(Y) \subseteq \mathfrak{g}_0(X) \tag{3.65}$$

because by construction $\mathrm{ad}(Y)$ cannot be nilpotent on the other spaces $\mathfrak{g}_\lambda(X)$. But by hypothesis the element X is regular, thus the inclusion (3.65) cannot be strict. Thus $\mathfrak{g}_0(X) \subset \mathfrak{g}_0(Y)$ which means that $\mathrm{ad}(Y)$ is nilpotent on $\mathfrak{g}_0(X)$.

Now the fact for $\mathrm{ad}(Y)$ to be nilpotent means the vanishing of a polynomial determined by the coefficients of the matrix of $\mathrm{ad}(Y)$. Since this polynomial vanishes on the open set $\mathcal{U}$, it vanishes identically, so that $\mathrm{ad}(Y)$ is nilpotent on $\mathfrak{g}_0(X)$. It results that $\mathfrak{g}_0(X)$ is a ad-nilpotent algebra and the Engel's theorem 3.32 concludes that $\mathfrak{g}_0(X)$ is nilpotent.

We still have to prove that $\mathfrak{g}_0(X)$ is its own centralizer. Since $\mathfrak{g}_0(X)$ is a subalgebra we have the inclusion

$$\mathfrak{g}_0(X) \subseteq \mathcal{Z}\big(\mathfrak{g}_0(X)\big). \tag{3.66}$$

Let $Z \in \mathcal{Z}\big(\mathfrak{g}_0(X)\big)$. For each $Y \in \mathfrak{g}_0(X)$ we have $[Z,Y] \in \mathfrak{g}_0(X)$. In particular with $Y = X$ we have $\mathrm{ad}(X)Z \in \mathfrak{g}_0(X)$. Thus

$$\mathrm{ad}(X)^n Z = \mathrm{ad}(X)^{n-1}\underbrace{\mathrm{ad}(X)Z}_{\in\mathfrak{g}_0(X)} \tag{3.67}$$

and there exists a n such that $\mathrm{ad}(X)^{n-1}\,\mathrm{ad}(X)Z = 0$. □

If $\mathfrak{g}$ is a Lie algebra, the group of **inner automorphism** is the subgroup of $\mathrm{Aut}(\mathfrak{g})$ generated by the elements of the form $e^{\mathrm{ad}(X)}$ with $X \in \mathfrak{g}$.

Theorem 3.72.
The group of inner automorphisms of $\mathfrak{g}$ acts transitively on the set of Cartan subalgebras.

For a proof, see [18]. In particular they have all the same dimension and the definition of the **rank** as the dimension of its Cartan algebra make sense. In [18] we have a more abstract definition of the rank, see page III-2.

Proposition 3.73.
If $\mathfrak{h}$ is a Cartan subalgebra of the complex Lie algebra $\mathfrak{g}$, there exists a regular element $X \in \mathfrak{g}$ such that $\mathfrak{h} = \mathfrak{g}_0(X)$.

For a proof, see [18].

[14] this is made more explicitly in the proof of theorem 3.77.
[15] it means that $\mathrm{ad}(Y)$ is invertible.

Proposition 3.74.
A Cartan subalgebra is a maximal nilpotent subalgebra.

Proof. Let $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$ and $\mathfrak{n}$, a nilpotent algebra which contains $\mathfrak{h}$. Let $\{X_1, \ldots, X_n\}$ be a basis of $\mathfrak{g}$ chosen in such a way that the p first vectors form a basis of $\mathfrak{h}$ while the r first, a basis of $\mathfrak{n}$ ($r > p$ of course). As notational convention, the subscript i, j are related to $\mathfrak{h}$ and u, t to $\mathfrak{n} \ominus \mathfrak{h}$.

Let us first suppose $\dim \mathfrak{n} = \dim \mathfrak{h} + 1$ and let X_u be the basis vector of $\mathfrak{n}$ which is not in $\mathfrak{h}$. Since $\mathfrak{h}$ is Cartan, we can find $X_i \in \mathfrak{h}$ such that $Y = [X_u, X_i] \notin \mathfrak{h}$. Then Y has a X_u-component and this contradict the fact that $\operatorname{ad} X_i$ is nilpotent.

The next case is $\mathfrak{n} = \mathfrak{h} \oplus X_u \oplus X_t$. In this case we can find a $X_i \in \mathfrak{h}$ such that $Y = [X_u, X_i] \notin \mathfrak{h}$. The fact to be nilpotent makes that Y has no X_u-component, so that it has a X_t-component. Now it is clear that for any $X_j \in \mathfrak{h}$, $[Y, X_j]$ still has no X_u-component (because $(\operatorname{ad} X_i \circ \operatorname{ad} X_j)$ has to be nilpotent), but has also no X_t-component. Then for any $X \in \mathfrak{h}$, $[Y, X] \in \mathfrak{h}$ with $Y \notin \mathfrak{h}$. There is a contradiction.

Now the step to the general case is easy: if $\dim \mathfrak{n} = \dim \mathfrak{h} + m$, we consider $X_1, \ldots, X_m \in \mathfrak{h}$ and $A = (\operatorname{ad} X_1 \circ \operatorname{ad} X_m) X_u$. This is not in $\mathfrak{h}$ although $[A, X] \in \mathfrak{h}$ for any $X \in \mathfrak{h}$. □

Proposition 3.75.
If $\mathfrak{g}$ is a semisimple Lie algebra, a subalgebra $\mathfrak{h}$ is Cartan if and only if the two following conditions are satisfied:

(i) $\mathfrak{h}$ *is a maximal abelian subalgebra*

(ii) the endomorphism $\operatorname{ad}(H)$ *is diagonalizable for every* $H \in \mathfrak{h}$.

3.8 Root spaces in semisimple complex Lie algebras

In this section we particularize ourself to complex semisimple Lie algebras. A very good reference about complex semisimple algebras including the reconstruction *via* the Cartan matrix and Chevalley-Weyl basis is [18].

3.8.1 Introduction and notations

Real and complex Lia algebras deserve quite different treatment with root space. We review here the main steps in both cases, emphasising the differences. We restrict ourself to semisimple Lie algebras. See [16].

3.8.1.1 Complex Lie algebras

If $\mathfrak{g}$ is a complex semisimple Lie algebra, we choose a Cartan subalgebra $\mathfrak{h}$ and the root spaces are given by

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \text{ st } [H, X] = \alpha(H) X \forall H \in \mathfrak{h}\}. \tag{3.68}$$

The dimension of $\mathfrak{h}$ is the rank of $\mathfrak{g}$. Then the root space decomposition reads

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha \tag{3.69}$$

where Φ is the set of roots.

3.8.1.2 Real Lie algebras

If $\mathfrak{g}$ is a real semisimple Lie algebra we consider a Cartan involution and the Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$. Then we choose a maximally abelian subalgebra $\mathfrak{a}$ in $\mathfrak{p}$ and we define

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} \text{ st } [J, X] = \alpha(J) X \forall J \in \mathfrak{a}\}. \tag{3.70}$$

The rank of $\mathfrak{g}$ is the dimension of $\mathfrak{a}$. The root space decomposition then reads

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\lambda \in \Sigma} \mathfrak{g}_\lambda \tag{3.71}$$

where Σ is the set of $\lambda \in \mathfrak{a}^*$ such that $\lambda \neq 0$ and $\mathfrak{g}_\lambda \neq 0$.

3.8.1.3 Notations

We summarize the notations that will be used later. Let $\mathfrak{h}$ be a Cartan algebra in the complex semisimple Lie algebra $\mathfrak{g}$. An element $\alpha \in \mathfrak{h}^*$ is a root if the space

$$\mathfrak{g}_\alpha = \{X \in \mathfrak{g} \text{ st } \operatorname{ad}(H)X = \alpha(H)x, \forall H \in \mathfrak{h}\} \tag{3.72}$$

is non empty.

(i) Φ is the set of all the roots. We consider an ordering notion on Φ and $\Phi^+ = \Pi$ is the set of positive roots.

(ii) An element in Φ^+ is simple if it cannot be written as the sum of two positive roots.

(iii) Δ is the set of simple roots[16]. The simple roots are denoted by $\{\alpha_1, \ldots, \alpha_l\}$.

3.8.2 Root spaces

We are considering a complex semisimple Lie algebra $\mathfrak{g}$ with a Cartan subalgebra $\mathfrak{h}$.

Definition 3.76.
For each $\alpha \in \mathfrak{h}^*$ we define

$$\mathfrak{g}_\alpha = \{x \in \mathfrak{g} \textit{ st } \forall h \in \mathfrak{h}, \big(\operatorname{ad} h - \alpha(h)\big)^n x = 0 \textit{ for some } n \in \mathbb{N}\}. \tag{3.73}$$

*If $\mathfrak{g}_\alpha$ is not reduced to 0, we say that α is a **root** and $\mathfrak{g}_\alpha$ is a **root space**.*

Corollary 3.84 will provide an easier formula for the root spaces when the algebra $\mathfrak{g}$ is complex and semisimple.

Theorem 3.77.
Let $\mathfrak{g}$ be a complex Lie algebra with Cartan subalgebra $\mathfrak{h}$. If $\alpha, \beta \in \mathfrak{g}^*$ then

(i) $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$*,*

(ii) $\mathfrak{g}_0 = \mathfrak{h}$*.*

Proof. For $z \in \mathfrak{h}$ and x, $y \in \mathfrak{g}$ we have

$$\big(\operatorname{ad} z - (\alpha+\beta)(z)\big)[x,y] = [(\operatorname{ad} z - \alpha(z))x, y] + [x, (\operatorname{ad} z - \beta(z))y]. \tag{3.74}$$

Now suppose that for some n,

$$\big(\operatorname{ad} z - (\alpha+\beta)(z)\big)^n [x,y] = \sum_k \binom{k}{n}\binom{k}{n} \left[(\operatorname{ad} z - \alpha(z))^k(x), (\operatorname{ad} z - \beta(z))^{n-k}(y)\right]. \tag{3.75}$$

If we apply $(\operatorname{ad} z - (\beta+\alpha)(z))^n$ to this equality, we find

$$\begin{aligned}
&(\operatorname{ad} z - (\beta+\alpha)(z))^{n+1}[x,y] \\
&\quad = \sum_{k=1}^{n} \binom{k}{n} \Big(\big[(\operatorname{ad} z - \alpha(z))(\operatorname{ad} z - \alpha(z))^k(x), (\operatorname{ad} z - \beta(z))^{n-k}(y)\big] \\
&\qquad + \big[(\operatorname{ad} z - \alpha(z))(\operatorname{ad} z - \alpha(z))^k(x), (\operatorname{ad} z - \beta(z))^{n-k+1}(y)\big]\Big) \\
&\quad = \sum_{k=1}^{n+1} \binom{k}{n+1} \big[(\operatorname{ad} z - \alpha(z))^k(x), (\operatorname{ad} z - \beta(z))^{n+1-k}(y)\big].
\end{aligned} \tag{3.76}$$

This formula shows that $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$. Indeed let $x \in \mathfrak{g}_\alpha$, $y \in \mathfrak{g}_\beta$ and n be large enough,

$$(\operatorname{ad} z - (\alpha+\beta)(z))^n [x,y] = 0. \tag{3.77}$$

Now we turn our attention to the second part. Let us apply the Lie theorem 3.20 to the action of $\mathfrak{g}$ on the quotient $\mathfrak{g}_0/\mathfrak{h}$. There exists $[X_0] \in \mathfrak{g}_0/\mathfrak{h}$ such that $h[X_0] = \lambda(h)[X_0]$ where the bracket stand for the class. Since $\mathfrak{h}$ is nilpotent on $\mathfrak{g}_0$ we have $\lambda = 0$ identically. Looking outside the class, the existence of a non vanishing $[X_0] \in \mathfrak{g}/\mathfrak{h}$ such that $h[X_0] = 0$ means that there exists $X_0 \in \mathfrak{g}_0 \backslash \mathfrak{h}$ such that $[h, X_0] \in \mathfrak{h}$ for every $h \in \mathfrak{h}$. This contradicts the fact that $\mathfrak{h}$ is its own centralizer. $\square$

[16]The symbol Δ has not a fixed signification in the literature. As example, in [19] the symbol Δ is the set of roots while in [9] it denotes the set of simple roots.

Proposition 3.78.
The complex Lie algebra decomposes into the root spaces as

$$\mathfrak{g} = \bigoplus_{\alpha\in\mathfrak{h}^*} \mathfrak{g}_\alpha. \tag{3.78}$$

Proof. Let $H \in \mathfrak{h}$. We consider the primary decomposition (3.60) with respect to the operator $\operatorname{ad}(H)$:

$$\mathfrak{g} = \bigoplus_\lambda \mathfrak{g}_\lambda(H). \tag{3.79}$$

If $H' \in \mathfrak{h}$ the operator $\operatorname{ad}(H')$ acts the space $\mathfrak{g}_\lambda(H)$ because $H' \in \mathfrak{g}_0(H)$ so that

$$[H', \mathfrak{g}_\lambda(H)] \subset \mathfrak{g}_\lambda(H). \tag{3.80}$$

Thus we can write the primary decomposition of $\mathfrak{g}_\lambda(H)$ with respect to the operator $\operatorname{ad}(H')$ knowing that

$$\big(\mathfrak{g}_\lambda(H)\big)_\mu(H') = \{X \in \mathfrak{g}_\lambda(H) \text{ st } \big(\operatorname{ad}(H') - \mu\big)^n X = 0\} = \mathfrak{g}_\lambda(H) \cap \mathfrak{g}_\mu(H'). \tag{3.81}$$

What we get is the decomposition

$$\mathfrak{g} = \bigoplus_\lambda \bigoplus_\mu \mathfrak{g}_\lambda(H) \cap \mathfrak{g}_\mu(H'). \tag{3.82}$$

We continue the decomposition with $H'', H''', \ldots$ until each $\operatorname{ad}(H)$ with $H \in \mathfrak{h}$ has only one eigenvalue on each of the summand of the decomposition

$$\mathfrak{g} = \bigoplus_{\lambda_1,\ldots,\lambda_l} \mathfrak{g}_{\lambda_1}(H_1) \cap \ldots \cap \mathfrak{g}_{\lambda_l}(H_l). \tag{3.83}$$

For each l-uple $(\lambda_1, \ldots, \lambda_l)$, the eigenvalue of H_i on $\mathfrak{g}_{\lambda_1} \cap \ldots \cap \mathfrak{g}_{\lambda_l}$ is λ_i. Thus we can see λ as a 1-form on $\mathfrak{h}$ and write

$$\mathfrak{g} = \bigoplus_\lambda \mathfrak{g}_\lambda \tag{3.84}$$

with

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} \text{ st } \big(\operatorname{ad}(H) - \lambda(H)\big)^n X = 0\}. \tag{3.85}$$

□

Corollary 3.79.
If $X_\alpha \in \mathfrak{g}_\alpha$ and $X_\beta \in \mathfrak{g}_\beta$ with $\alpha + \beta \neq 0$, then $B(X_\alpha, X_\beta) = 0$.

Proof. From the second point of proposition 3.77, we have $\operatorname{ad} X_\alpha \circ \operatorname{ad} X_\beta \colon \mathfrak{g}_\mu \to \mathfrak{g}_{\mu+\alpha+\beta}$. If $\alpha + \beta \neq 0$, the fact that the sum (3.84) is direct makes the trace of $\operatorname{ad} X_\alpha \circ \operatorname{ad} X_\beta$ zero. □

Since $\mathfrak{g}$ is semisimple, the restriction of the Killing form on $\mathfrak{h}$ is nondegenerate[17]. Thus we can introduce, for each linear function $\phi \colon \mathfrak{h} \to \mathbb{C}$, the unique element $t_\phi \in \mathfrak{h}$ such that

$$\phi(h) = B(t_\phi, h) \tag{3.86}$$

for every $h \in \mathfrak{h}$. This element is nothing else that the dual ϕ^* with respect to the Killing form. Indeed

$$t^*_\phi(h) = B(t_\phi, h) = \phi(h), \tag{3.87}$$

so that $t^*_\phi = \phi$. Incidentally, this proves that when ϕ runs over a basis of $\mathfrak{h}^*$, the vector t_ϕ runs over a basis of $\mathfrak{h}$. The space $\mathfrak{h}^*$ is endowed with an inner product defined by

$$(\alpha, \beta) = B(t_\alpha, t_\beta) = \beta(t_\alpha) = \alpha(t_\beta). \tag{3.88}$$

Lemma 3.80.
If $X \in \mathfrak{g}_\alpha$ and $Y \in \mathfrak{g}_{-\alpha}$, then

$$[X, Y] = B(X, Y) t_\alpha. \tag{3.89}$$

Proof. By theorem 3.77(i), $[X, Y] \in \mathfrak{g}_0 = \mathfrak{h}$. Now we consider $h \in \mathfrak{h}$ and the invariance formula (3.21). We find:

$$B\big(h, [X, Y]\big) = -B\big([X, h], Y\big) = \alpha(h) B(X, Y) = B(h, t_\alpha) B(X, Y) = B\big(h, B(X, Y) t_\alpha\big). \tag{3.90}$$

The lemma is proven since it is true for any $h \in \mathfrak{h}$ and B is nondegenerate on $\mathfrak{h}$. □

[17] Because the Killing form is zero on each space $\mathfrak{g}_\alpha$ with $\alpha \neq 0$.

The elements t_α allow to introduce an inner product on $\mathfrak{h}^*$ and hence on the roots by defining

$$(\alpha, \beta) = B(t_\alpha, t_\beta). \tag{3.91}$$

Lemma 3.81.
If α and β are roots we have the formula

$$(\alpha, \beta) = \sum_{\gamma \in \Phi} (\dim \mathfrak{g}_\gamma)(\alpha, \gamma)(\beta, \gamma). \tag{3.92}$$

Proof. We consider for $\mathfrak{g}$ a basis in which all the elements are part of one of the root spaces and we look at the endomorphism $\mathrm{ad}(t_\alpha)$ of $\mathfrak{g}$. This is diagonal and has zeros on the entries corresponding to $\mathfrak{h}$. The other entries on the diagonal are of the form $\gamma(t_\alpha)$. Thus

$$B(t_\alpha, t_\beta) = \sum_{\gamma \in \Phi} (\dim \mathfrak{g}_\gamma)\gamma(t_\alpha)\gamma(t_\beta). \tag{3.93}$$

Thus we have $(\alpha, \beta) = B(t_\alpha, t_\beta) = \sum_{\gamma \in \Phi} (\dim \mathfrak{g}_\gamma)(\alpha, \gamma)(\beta, \gamma)$. □

Proposition 3.82.
Let α and β be roots. We have

(i) $(\alpha, \beta) \in \mathbb{Q}$,

(ii) $(\alpha, \alpha) \geqslant 0$.

The proof comes from [19] page 826.

Proof. Let $\alpha, \beta \in \Phi$ and consider the space

$$V = \bigoplus_{m \in \mathbb{Z}} \mathfrak{g}_{\beta + m\alpha}. \tag{3.94}$$

If $X_\alpha \in \mathfrak{g}_\alpha$ and $X_{-\alpha} \in \mathfrak{g}_{-\alpha}$ with $[X_\alpha, X_{-\alpha}] = t_\alpha$ we have, for all $v \in V$,

$$[X_\alpha, v] \in V \tag{3.95a}$$
$$[X_{-\alpha}, v] \in V \tag{3.95b}$$
$$[t_\alpha, v] \in V. \tag{3.95c}$$

Thus we can consider the restrictions to V of the operators $\mathrm{ad}(X_\alpha)$, $\mathrm{ad}(X_{-\alpha})$ and $\mathrm{ad}(t_\alpha)$. Since ad is an homomorphism we have, as operator on V,

$$\mathrm{ad}(t_\alpha) = \big[\mathrm{ad}(X_\alpha), \mathrm{ad}(X_{-\alpha})\big], \tag{3.96}$$

and then $\mathrm{Tr}\big(\mathrm{ad}(t_\alpha)|_V\big) = 0$.

Let us compute that trace on the basis $\{v_k^{(i)}\}$ where $v_k^{(i)} \in \mathfrak{g}_{\beta + k\alpha}$. Since

$$\mathrm{ad}(t_\alpha)v_k^{(i)} = (\beta + k\alpha)(t_\alpha)v_k^{(i)} \tag{3.97}$$

we have

$$0 = \mathrm{Tr}\big(\mathrm{ad}(t_\alpha)|_V\big) \tag{3.98a}$$
$$= \sum_{k \in \mathbb{Z}} \dim \mathfrak{g}_{\beta + k\alpha}(\beta + k\alpha)(t_\alpha) \tag{3.98b}$$
$$= \sum_{k \in \mathbb{Z}} \dim_{\beta + k\alpha}\big((\alpha, \beta) + (\alpha, \alpha)\big) \tag{3.98c}$$

and

$$\underbrace{\Big(\sum_{k \in \mathbb{Z}} \dim \mathfrak{g}_{\beta + k\alpha}\Big)}_{A \in \mathbb{N}}(\alpha, \beta) = -(\alpha, \alpha)\underbrace{\Big(\sum_{k \in \mathbb{Z}} k \dim \mathfrak{g}_{\beta + k\alpha}\Big)}_{B \in \mathbb{Z}}. \tag{3.99}$$

If $(\alpha, \alpha) = 0$ then we have $(\beta, \alpha) = 0$ for every $\beta \in \Phi$, hence $B(t_\alpha, t_\beta) = 0$ which contradicts non degeneracy of the Killing form. We conclude that $(\alpha, \alpha) \neq 0$. By the formula of lemma 3.81 we get

$$(\alpha, \alpha) = \sum_{\beta \in \Phi} \dim \mathfrak{g}_\beta (\alpha, \beta)^2. \tag{3.100}$$

Replacing in that formula the value of (α, β) taken from formula (3.99) we found

$$(\alpha, \alpha) = \sum_{\beta \in \Phi} \dim \mathfrak{g}_\beta \frac{B^2}{A^2} (\alpha, \alpha)^2 \tag{3.101}$$

and then $(\alpha, \alpha) \in \mathbb{Q}^+$. The fact that (α, β) is rational follows.

Notice that the sign of B is not guaranteed because it's not sure because we do not know whether there are more positive or negative terms in the sum of the right hand side of (3.99). □

Proposition 3.83.
Let α be a root of the complex semisimple Lie algebra $\mathfrak{g}$. Then

(i) $\dim \mathfrak{g}_\alpha = 1$,

(ii) the only integer multiple of α to be roots are $\pm\alpha$.

Proof. Let $X_\alpha \in \mathfrak{g}_\alpha$ and consider the vector space

$$V = \mathbb{C} t_\alpha \oplus \mathbb{C} X_\alpha \oplus \bigoplus_{m<0} \mathfrak{g}_{m\alpha}. \tag{3.102}$$

Let $y \in \mathfrak{g}_{-\alpha}$ be chosen in such a way that $[X_\alpha, y] = t_\alpha$; by lemma 3.80 this is only a matter of normalization. The space V is invariant under $\mathrm{ad}(X_\alpha)$ and $\mathrm{ad}(y)$. Indeed

(i) $\mathrm{ad}(X_\alpha) t_\alpha = -\alpha(t_\alpha) X_\alpha \in \mathbb{C} X_\alpha$;

(ii) $\mathrm{ad}(X_\alpha) X_\alpha = 0$;

(iii) $\mathrm{ad}(X_\alpha) \mathfrak{g}_{m\alpha} \subset \mathfrak{g}_{(m+1)\alpha}$; if $m < -1$, $(m+1) < 0$, while if $m = -1$ we know that the commutator $[X_\alpha, \mathfrak{g}_{-\alpha}]$ is included in $\mathbb{C} t_\alpha \in V$;

(iv) $\mathrm{ad}(y) t_\alpha \in \mathfrak{g}_{-\alpha}$

(v) $\mathrm{ad}(y) X_\alpha = -t_\alpha$ by definition;

(vi) $\mathrm{ad}(y) \mathfrak{g}_{m\alpha} \subset \mathfrak{g}_{(m-1)\alpha}$.

Since $\mathrm{ad}\colon \mathfrak{g} \to \mathrm{GL}(\mathfrak{g})$ is an homomorphism (lemma 3.4) we have

$$\big[\mathrm{ad}(X_\alpha), \mathrm{ad}(y) \big] = \mathrm{ad}(t_\alpha) \tag{3.103}$$

and then $\mathrm{Tr}\big(\mathrm{ad}(t_\alpha)\big) = 0$ because the trace of a commutator is zero[18]. Since V is an invariant subspace, the trace of $\mathrm{ad}(t_\alpha)$ restricted to V is also vanishing. Let us compute that trace on the basis $\{X_\alpha, t_\alpha, X^i_{m\alpha}\}_{m<0}$ where i takes as many values as the dimension of $\mathfrak{g}_{m\alpha}$.

We have $\mathrm{ad}(t_\alpha) X_{-\alpha} = -\alpha(t_\alpha) X_{-\alpha}$, $\mathrm{ad}(t_\alpha) t_\alpha = 0$ and $\mathrm{ad}(t_\alpha) X^i_{m\alpha} = m\alpha(t_\alpha) X_{m\alpha}$, thus the trace is

$$0 = \alpha(t_\alpha) \Big(-1 + \sum_{m=1}^{\infty} m \dim \mathfrak{g}_{m_\alpha} \Big). \tag{3.104}$$

Notice that the sum is in fact finite since the dimension of $\mathfrak{g}$ is finite. We know that $\alpha(t_\alpha) = B(t_\alpha, t_\alpha) \neq 0$, so that equation (3.104) is only possible with $\dim \mathfrak{g}_\alpha = 1$ and $\dim \mathfrak{g}_{m\alpha} = 0$ for $m \neq 1$. □

A very similar proof can be found in [19], page 827.

Corollary 3.84.
In the case of semisimple complex Lie algebra,

(i) the root spaces are given by

$$\mathfrak{g}_\alpha = \{ X \in \mathfrak{g} \text{ st } \forall h \in \mathfrak{h}, [h, X] = \alpha(h) X \}; \tag{3.105}$$

(ii) for every $x_\alpha \in \mathfrak{g}_\alpha$, and for every $h \in \mathfrak{h}$, we have

$$[h, x_\alpha] = \alpha(h) x_\alpha. \tag{3.106}$$

[18] From the cyclic invariance of the trace.

Proof. Let $X \in \mathfrak{g}_\alpha$, we have

$$\big(\operatorname{ad}(h) - \alpha(h)\big)^n X = 0, \tag{3.107}$$

so

$$\big(\operatorname{ad}(h) - \alpha(h)\big) \underbrace{\big(\operatorname{ad}(h) - \alpha(h)\big)^{n-1} X}_{v} = 0. \tag{3.108}$$

In particular the vector $v = \big(\operatorname{ad}(h) - \alpha(h)\big)^{n-1} X$ belongs to $\mathfrak{g}_\alpha$. Since the latter space has dimension one, the vector v is a multiple of X and consequently equation (3.108) shows that

$$\big(\operatorname{ad}(h) - \alpha(h)\big) v = \big(\operatorname{ad}(h) - \alpha(h)\big) X = 0. \tag{3.109}$$

The second point is only an other way to write the same. □

Lemma 3.85.
If H is an element of $\mathfrak{h}$ with $\alpha(H) = 0$ for every root, then $H = 0$

Proof. Consider the decompositions (not unique) $H = \sum_{\alpha\in\Phi} a_\alpha t_\alpha$ and $H' = \sum_{\beta\in\phi} a'_\beta t_\beta$. Then

$$B(H, H') = \sum_{\alpha,\beta} a_\alpha a'_\beta B(t_\alpha, t_\beta) \tag{3.110a}$$
$$= \sum_{\alpha,\beta} a'_\beta \beta(a_\alpha, t_\alpha) \tag{3.110b}$$
$$= \sum_{\beta} a'_\beta \beta(H) \tag{3.110c}$$
$$= 0. \tag{3.110d}$$

Such an element is thus Killing-orthogonal to the whole space $\mathfrak{h}$ but we already know the $\mathfrak{h}$ is orthogonal to each space $\mathfrak{g}_\alpha$ ($\alpha \neq 0$). By non degeneracy of the Killing form we must have $H = 0$. □

Proposition 3.86.
The set of roots of a complex semisimple Lie algebra spans the dual space $\mathfrak{h}^*$.

Proof. Consider a basis $\{H_i\}$ of $\mathfrak{h}$ with $\{H_0, \ldots, H_m\} = \operatorname{Span}(\Phi)$ and $\{H_{m+1}, \ldots, H_r\}$ be outside of $\operatorname{Span}\Phi$. A root reads $\alpha = \sum_{k=0}^m a_k H_k^*$. Thus $\alpha(H_{m+1}) = 0$, which implies that $H_{m+1} = 0$ by lemma 3.85. □

Corollary 3.87.
A Cartan algebra $\mathfrak{h}$ of a complex semisimple Lie algebra is abelian.

Proof. Let $H', H'' \in \mathfrak{h}$ and consider $H = [H', H'']$, a root α and $X_\alpha \in \mathfrak{g}_\alpha$. On the one hand we have

$$\big[[H', H''], X_\alpha\big] = -\alpha(H')[X_\alpha, H'] + \alpha(H')[X_\alpha, H''] = 0 \tag{3.111}$$

and on the other hand we have $\big[[H', H''], X_\alpha\big] = [H, X_\alpha] = \alpha(H) X_\alpha$. We deduce that $\alpha(H) = 0$ for every root and then that $H = 0$ by lemma 3.85. □

We denote by Φ the set of roots. These are the elements $\lambda \in \mathfrak{h}^*$ such that $\mathfrak{g}_\lambda$ is non trivial. We suppose to have chosen a positivity notion on $\mathfrak{h}^*$, so that we can speak of Φ^+, the set of **positive roots**.

A positive root is **simple** is it cannot be written as the sum of two positive roots.

3.8.3 Generators

We are going to build the Chevalley basis of the complex semisimple Lie algebra $\mathfrak{g}$. That will essentially be a choice of a basis vector in each of the root spaces. We are following the notations summarized in point 3.8.1.3.

Now, for each root α, we pick $e_\alpha \in \mathfrak{g}_\alpha$. We will see that, up to renormalization, we can set the in nice commutation relations.

Lemma 3.88.
If α and β are roots such that $\alpha + \beta \neq 0$, then

$$B(e_\alpha, e_\beta) = 0. \tag{3.112}$$

If $f_\alpha \in \mathfrak{g}_{-\alpha}$ we also have $B(e_\alpha, f_\alpha) \neq 0$.

Proof. By definition $B(e_\alpha, e_\beta) = \operatorname{Tr}\big(\operatorname{ad}(e_\alpha)\circ\operatorname{ad}(e_\beta)\big)$. If we apply $\operatorname{ad}(e_\alpha)\circ\operatorname{ad}(e_\beta)$ to an element of e_γ (including $\mathfrak{g}_0 = \mathfrak{h}$), we get an element of $\mathfrak{g}_{\alpha+\beta+\gamma}$. Thus the trace defining the Killing form is zero and $B(e_\alpha, e_\beta) = 0$ when $\alpha + \beta = 0$.

Since the Killing form is nondegenerate, we conclude that $B(e_\alpha, e_{-\alpha}) \neq 0$. □

Corollary 3.89.
Let $\mathfrak{g}$ be a semisimple complex Lie algebra and $\mathfrak{h}$ be a Cartan subalgebra of $\mathfrak{g}$. Let α be a root of $\mathfrak{g}$ and $\mathfrak{h}_\alpha = [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}]$. There exist an unique $H_\alpha \in \mathfrak{h}_\alpha$ such that $\alpha(H_\alpha) = 2$.

Proof. We have $[e_\alpha, f_\alpha] = B(e_\alpha, f_\alpha)t_\alpha$ and the lemma 3.88 shows that the Killing form is non zero. Multiplying by a suitable number provides the result. □

The element $H_\alpha \in \mathfrak{h}$ defined in this lemma is the **inverse root** of α.

Lemma 3.90.
Let $\{\beta_1, \ldots, \beta_l\}$ be a choice of elements in $\mathfrak{h}^*$ such that the set $\{t_{\beta_1}, \ldots, t_{\beta_l}\}$ is a basis of $\mathfrak{h}$. Thus the roots can be decomposed as

$$\alpha = \sum_{k=1}^{l} a_k \beta_k \tag{3.113}$$

with $a_k \in \mathbb{Q}$.

Proof. Let $\alpha = \sum_{k=1}^{l} a_k\beta_k$. We know that the vectors t_{β_i} form a basis of $\mathfrak{h}$, so we have the decomposition $t_\alpha = \sum_k a_k t_{\beta_k}$. Indeed

$$B\big(h, \sum_k a_k t_{\beta_k}\big) = \sum_k a_k B(h, t_{\beta_k}) = \sum_k a_k \beta_k(h) = \alpha(h). \tag{3.114}$$

For each $k = 1, 2, \ldots, l$ we have

$$(\alpha_k, \alpha) = \sum_{j=1}^{l} a_k(\alpha_k, \alpha_j). \tag{3.115}$$

This is a system of linear equations for the l variables a_k. Since the coefficients (α_k, α) and (α_k, α_j) are rational by proposition 3.82, the solutions are rational too. □

Remark 3.91.
The lemma 3.90 deals with a quite general basis of $\mathfrak{h}$. We will see in the proposition 3.101 that in the case of the basis of simple roots, the coefficients a_k are integers, either all positive or all negative.

3.8.4 Subalgebra $\mathfrak{sl}(2)_i$

For each nonzero root $\alpha \in \mathfrak{h}^*$, we choose $e_\alpha \in \mathfrak{g}_\alpha$ and $f_\alpha \in \mathfrak{g}_{-\alpha}$ in such a way to have

$$B(e_\alpha, f_\alpha) = \frac{2}{B(t_\alpha, t_\alpha)}, \tag{3.116}$$

and then we pose

$$h_\alpha = \frac{2}{B(t_\alpha, t_\alpha)} t_\alpha. \tag{3.117}$$

Notice that these choices are possible because the Killing form is non degenerated on $\mathfrak{h}$.

The following comes from page 82 of [9]

Proposition 3.92.
For each root, the set $\{e_\alpha, f_\alpha, h_\alpha\}$ generates an algebra isomorphic to $\mathfrak{sl}(2, \mathbb{R})$, i.e. they satisfy

$$[h_\alpha, e_\alpha] = 2e_\alpha \tag{3.118a}$$
$$[h_\alpha, f_\alpha] = -2f_\alpha \tag{3.118b}$$
$$[e_\alpha, f_\alpha] = h_\alpha \tag{3.118c}$$
(3.118d)

Proof. Since $\alpha(t_\alpha) = B(t_\alpha, t_\alpha)$ we have $\alpha(h_\alpha) = 2$. Now the computations are quite direct. The first is

$$[h_\alpha, e_\alpha] = \alpha(h_\alpha)e_\alpha = 2e_\alpha. \tag{3.119}$$

For the second,

$$[h_\alpha, f_\alpha] = -\alpha(h_\alpha)f_\alpha = -2f_\alpha. \tag{3.120}$$

For the third, we know that $[e_\alpha, f_\alpha] \in \mathfrak{h}$; thus $B\big(X, [e_\alpha, f_\alpha]\big) = 0$ whenever $X \in \mathfrak{g}_\lambda$ with $\lambda \neq 0$. Let $h \in \mathfrak{h}$. Using the invariance of the Killing form,

$$B\big(h, [e_\alpha, f_\alpha]\big) = B\big([h, e_\alpha], f_\alpha\big) = \alpha(h)B(e_\alpha, f_\alpha) = B(t_\alpha, t_\alpha)B(e_\alpha, f_\alpha) = B\big(B(e_\alpha, f_\alpha)t_\alpha, h\big). \tag{3.121}$$

Thus

$$[e_\alpha, f_\alpha] = B(e_\alpha, f_\alpha)f_\alpha = h_\alpha. \tag{3.122}$$

□

Remark that we used the non degeneracy of the Killing form in a crucial way. The copy of $\mathfrak{sl}(2,\mathbb{R})$ formed by $\{e_\alpha, f_\alpha, h_\alpha\}$ is denoted by $\mathfrak{sl}(2,\mathbb{R})_\alpha$.

Proposition 3.93.
In the universal enveloping algebra,

$$[h_j, f_i^{k+1}] = -(k+1)\alpha_i(h_j)f_i^{k+1} \tag{3.123}$$

as generalisation of the previous one.

Proof. We use an induction over k. Since $\operatorname{ad}(h_j)$ is a derivation in $\mathcal{U}(\mathfrak{g})$, the induction hypothesis and the definition relation $[h, f_i] = -\alpha_i(h)f_i$ with $h = h_i$, we have

$$\begin{aligned}
\operatorname{ad}(h_j)f_i^{k+1} &= \big(\operatorname{ad}(h_j)f_i^k\big)f_i + f_i^k \operatorname{ad}(h_j)f_i. \\
&= -k\alpha + i(h_j)f_i^k f_i - \alpha_i(h_j)f_i^{k+1} \\
&= -(k+1)\alpha_i(h_j)f_i^{k+1}.
\end{aligned} \tag{3.124}$$

□

Now the Lie algebra $\mathfrak{g}$ can be seen as a $\mathfrak{sl}(2,\mathbb{R})$-module. As an example, for each choice of $\beta \in \Phi$, the algebra $\mathfrak{sl}(2)_\alpha$ acts on the vector space

$$V = \bigoplus_{k\in\mathbb{Z}} \mathfrak{g}_{\beta+k\alpha}. \tag{3.125}$$

The vector space $\mathfrak{g}$ carries thus several representations of $\mathfrak{sl}(2)$; this fact will be used in a crucial way during the proof of proposition 3.99.

3.8.5 Chevalley basis

The Chevalley basis corresponds to an other choice of normalization of the element e_α, h_α. If we set

$$\begin{cases} H_\alpha = K_\alpha t_\alpha & (3.126a) \\ E_\alpha = N_\alpha e_\alpha & (3.126b) \end{cases}$$

with

$$\begin{aligned}
K_\alpha &= \frac{2}{(\alpha, \alpha)} \\
N_\alpha &= \sqrt{\frac{2}{B(e_\alpha, e_{-\alpha})(\alpha, \alpha)}},
\end{aligned} \tag{3.127}$$

then we have the **Chevalley relations**:

$$[E_\alpha, E_{-\alpha}] = H_\alpha \tag{3.128a}$$

$$[H_\alpha, E_\beta] = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}E_\beta \tag{3.128b}$$

$$[H_\alpha, H_\beta] = 0. \tag{3.128c}$$

The last relation is only the fact that the Cartan subalgebra $\mathfrak{h}$ is abelian.

If $\{\alpha_i\}_{i=1,\ldots,l}$ is the set of simple roots, we consider the notation $X_i^+ = E_{\alpha_i}$, $X_i^- = E_{-\alpha_i}$, $H_i = H_{\alpha_i}$ and we introduce the **Cartan matrix**

$$A_{ij} = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}. \tag{3.129}$$

Reduced to the simple roots the relations (3.128) become

$$\begin{aligned} [X_i^+, X_j^-] &= \delta_{ij} H_i \\ [H_i, X_j^\pm] &= \pm A_{ij} X_j^\pm \\ [H_i, H_j] &= 0. \end{aligned} \tag{3.130}$$

The first relation comes from the fact that $\alpha_i - \alpha_j$ is not a root when α_i and α_j are simple roots.

Remark 3.94.
The idea behind the Chevalley relations is that the algebra $\mathfrak{g}$ is generated by the elements $X_i^\pm$, H_i and the commutation relations (3.130). Even if these elements do not form a basis (while the elements E_α, H_α do), one can define a function on $\mathfrak{g}$ by giving its values on $X_i^\pm$ and H_i providing one has a canonical way to extend it on commutators.

*The definition **??** of standard cobracket works in this way.*

Remark 3.95.
Notice that these relations do not give the value of

$$[E_\alpha, E_\beta] = N_{\alpha,\beta} E_{\alpha+\beta} \tag{3.131}$$

when $\alpha + \beta$ is a root.

Problem and misunderstanding 7.
*It has to be possible to compute $N_{\alpha,\beta}$, but I do not know how. The answer is given in equation (3.358) but I don't know where I got them. Maybe there are some hints in [19] (Il faut ajouter Cornwell à la bibliographie et enlever le problème **??**).*

Problem and misunderstanding 8.
It seems that A_{ij} is the larger integer k such that $\alpha_i + k\alpha_j$ is a root. This is the justification of the other Serre's relations that read

$$\operatorname{ad}^{1-A_{ij}}(X_i^\pm) X_j^\pm = 0. \tag{3.132}$$

That relation has to be written with the Chevalley's ones.

One can choose the coefficients in a more scientific way[18]. Let α be a positive root, let H_α be the inverse root of α and $e_\alpha \in \mathfrak{g}_\alpha$. We have

$$[e_\alpha, e_\beta] = \begin{cases} N_{\alpha,\beta} e_{\alpha+\beta} & \text{if } \alpha+\beta \text{ is a root} \\ 0 & \text{if } \alpha+\beta \text{ is not a root.} \end{cases} \tag{3.133}$$

We are going to find a multiple E_α of e_α in such a way to have in the same time

$$[E_\alpha, E_{-\alpha}] = H_\alpha \tag{3.134a}$$

$$N_{\alpha,\beta} = -N_{-\alpha,-\beta}. \tag{3.134b}$$

Let σ be an involutive automorphism of $\mathfrak{g}$ such that $\sigma|_{\mathfrak{h}} : -\mathrm{id}$. First we have $\sigma(\mathfrak{g}_\alpha) = \mathfrak{g}_{-\alpha}$ because

$$[h, \sigma(e_\alpha)] = \sigma[\sigma(h), e_\alpha] = -\sigma\alpha(h) e_\alpha = -\alpha(h)\sigma(e_\alpha) \tag{3.135}$$

for every $h \in \mathfrak{h}$ and $e_\alpha \in \mathfrak{g}_\alpha$. From corollary 3.89 there exist a number a_α such that

$$[e_\alpha, \sigma(e_\alpha)] = a_\alpha H_\alpha. \tag{3.136}$$

We pose

$$E_\alpha = \frac{1}{\sqrt{-a_\alpha}} e_\alpha \tag{3.137a}$$

$$E_{-\alpha} = -\sigma(E_\alpha). \tag{3.137b}$$

With that choice we immediately have $[E_\alpha, E_{-\alpha}] = H_\alpha$. We also have $N_{\alpha,\beta} = -N_{-\alpha,-\beta}$; in order to see it, consider

$$[\sigma E_\alpha, \sigma E_\beta] = \sigma[E_\alpha, E_\beta] = N_{\alpha,\beta}\sigma(E_{\alpha+\beta}) = -N_{\alpha,\beta} E_{-\alpha-\beta}. \tag{3.138}$$

But the same is also equal to

$$[-E_{-\alpha}, -E_{-\beta}] = [E_{-\alpha}, E_{-\beta}] = N_{-\alpha,-\beta} E_{-\alpha,-\beta}. \tag{3.139}$$

Proposition 3.96.
With these choices we have

$$N_{\alpha,\beta} = \pm(p+1) \tag{3.140}$$

where p is the largest integer j such that $\beta - j\alpha$ is a root.

$\boxed{\textbf{Problem and misunderstanding}}$ **9.**
I don't know a proof of that, but [18] gives a reference.

From proposition 3.80 we know that $t_\alpha \in \mathfrak{h}_\alpha$, so that H_α is a multiple of H_α. The proportionality factor is easy to fix since

$$\begin{aligned} \alpha(H_\alpha) &= 2 && \text{definition of } H_\alpha \\ \alpha(t_\alpha) &= (\alpha,\alpha) && \text{definition (3.88)}. \end{aligned} \tag{3.141}$$

Thus $H_\alpha = \frac{2}{(\alpha,\alpha)} t_\alpha$ and

$$[H_\alpha, E_\beta] = \beta(H_\alpha)E_\beta = \frac{2}{(\alpha,\alpha)}\beta(t_\alpha) = \frac{2(\alpha,\beta)}{(\alpha,\beta)} \tag{3.142}$$

again by the definition (3.88).

3.8.6 Coefficients in the Cartan matrix

In this section we search to give the form of the coefficients in the Cartan matrix. We will show that the values of (α,β) are quite restricted.

Remark 3.97.
The notations are not standard. Here the symbol Δ denotes the set of simple *roots while the set of all roots is denoted by Φ. In the book [19], the symbol Δ is the set of all roots. This makes quite a difference !*

Definition 3.98.
*If α and β are roots of the complex semisimple Lie algebra $\mathfrak{g}$, then the α-**string** containing β is the set of roots of the form $\alpha + k\beta$ with $k \in \mathbb{Z}$.*

Among other things, the following proposition shows that a string has no gap.

Proposition 3.99.
Let $\alpha, \beta \in \Phi$. Then there exits integers p, q such that $\{\beta + k\alpha\}_{-p \leqslant k \leqslant q}$ is the α-string containing β. The numbers p and q satisfy

$$p - q = \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \tag{3.143}$$

and the form

$$\beta - \frac{2(\beta,\alpha)}{(\alpha,\alpha)}\alpha \tag{3.144}$$

is a nonzero root.

Proof. We consider the vector space

$$V = \bigoplus_{k\in\mathbb{Z}} \mathfrak{g}_{\beta+k\alpha} \tag{3.145}$$

and the Lie algebra $\mathfrak{sl}(2)_\alpha$ which acts on V. Simple computation using the fact that $\beta(h_\alpha) = 2(\alpha,\beta)/(\alpha,\alpha)$ shows that

$$[\frac{1}{2}h_\alpha, e_{\beta+k\alpha}] = \left(\frac{(\alpha,\beta)}{(\alpha,\alpha)} + k\right) e_{\beta+k\alpha}. \tag{3.146}$$

Thus the matrix of $\operatorname{ad}(\frac{1}{2}h_\alpha)$ is diagonal and has no multiplicity in its eigenvalues. We deduce that the representation if irreducible. From general theory of irreducible representations of $\mathfrak{sl}(2)$ we know that there exists an half-integer number j such that the diagonal entries of $\operatorname{ad}(\frac{1}{2}h_\alpha)$ take *all* the values from $-j$ to j by integer steps. Thus the α-string containing β has the form $\{\beta + k\alpha\}_{-p\leqslant k\leqslant q}$ where p and q satisfy

$$\begin{cases} \dfrac{(\alpha,\beta)}{(\alpha,\alpha)} - p = -j & \text{(3.147a)} \\ \dfrac{(\alpha,\beta)}{(\alpha,\alpha)} + q = j. & \text{(3.147b)} \end{cases}$$

Summing we get

$$p - q = \frac{2(\alpha, \beta)}{(\alpha, \alpha)}. \tag{3.148}$$

If λ is an eigenvalue of $\operatorname{ad}(\frac{1}{2}h_\alpha)$, then $-\lambda$ is also an eigenvalue (this is still from the irreducible representation theory of $\mathfrak{sl}(2)$). The number $(\alpha, \beta)/(\alpha, \beta)$ is obviously an eigenvalue (with $k = 0$), thus the string contains a k such that

$$\frac{(\alpha, \beta)}{(\alpha, \alpha)} + k = -\frac{(\alpha, \beta)}{(\alpha, \alpha)}. \tag{3.149}$$

The solution is $k = -2(\alpha, \beta)/(\alpha, \alpha)$ and we deduce that

$$\beta - 2\frac{(\alpha, \beta)}{(\alpha, \alpha)}\alpha \tag{3.150}$$

is a root of $\mathfrak{g}$. □

Proposition 3.100.
Let α, β be two roots. Then we have

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = 0, \pm 1, \pm 2, \pm 3. \tag{3.151}$$

Proof. First, equation (3.143) shows that $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ is integer. If $\alpha = \pm\beta$, the result is 2. If $\alpha \neq \pm\beta$, the vectors t_α and t_β are linearly independent and the Schwarz inequality shows

$$(\alpha, \beta)^2 = |B(t_\alpha, t_\beta)| < B(t_\alpha, t_\alpha)B(t_\beta, t_\beta) = (\alpha, \alpha)(\beta, \beta). \tag{3.152}$$

Thus

$$\left|\frac{2(\alpha, \beta)}{(\alpha, \alpha)}\right|\left|\frac{2(\alpha, \beta)}{(\beta, \beta)}\right| < \frac{4|(\alpha, \beta)(\alpha, \beta)|}{(\alpha, \beta)^2} = 4. \tag{3.153}$$

Consequently the number $|2(\alpha, \beta)/(\alpha, \alpha)|$ being integer can only take the values 0, 1, 2 and 3. Notice that the inequality in (3.152) and (3.153) are strict since α_i is not collinear to α_j. □

3.8.7 Simple roots

As seen before, Φ admits an order inherited from $\mathfrak{h}_{\mathbb{R}}^*$. A root $\alpha > 0$ is **simple** if it cannot be written as a sum of two positive roots.

Theorem 3.101.
Let $\{\alpha_1, \ldots, \alpha_l\}$ be the set of simple roots. Then every root $\beta \in \Phi$ can be decomposed into

$$\beta = \sum_{i=1}^{l} n_i \alpha_i \tag{3.154}$$

where non vanishing the numbers $n_i \in \mathbb{Z}$ are either all positive or all negative.

Proof. Let β be positive. If it is not simple, the one can decompose it into two positive roots:

$$\beta = \gamma + \delta \tag{3.155}$$

with $\gamma, \delta > 0$. If γ and/or δ are not simple, they can be decomposed further. This process has to be finite, indeed if the process is not finite, the decomposition of at least one positive root has to contains itself (because there are finitely many of them) while it is impossible to have $\gamma = \gamma + \alpha$ with $\alpha > 0$. □

Two corollaries: a root is either positive or negative (this is part of the definition of positivity) and when a root is positive, its decomposition into simple roots has only positive coefficients.

Lemma 3.102.
If $\alpha - \beta$ are simple roots with $\alpha \neq \beta$, then $\beta - \alpha$ is not a root and $B(h_\alpha, h_\beta) \leqslant 0$.

Proof. Define $\gamma = \beta - \alpha \in \Delta$ (and not Φ because $\alpha \neq \beta$). If $\gamma > 0$, the fact that $\beta = \gamma + \alpha$ contradict the simplicity of β while if $\gamma < 0$, in the same way $\alpha = \beta - \alpha$ contradict the simplicity of α.

Since $\beta - \alpha$ is not a root, $\beta_\alpha = 0$ and $\beta^\alpha \geqslant 0$ thus formula $2\beta(h_\alpha) = (\beta_\alpha - \beta^\alpha)\alpha(h_\alpha)$ gives

$$2B(h_\alpha, h_\beta) = \underbrace{(\beta_\alpha - \beta^\alpha)}_{\leqslant 0} B(h_\alpha, h_\alpha). \tag{3.156}$$

Now proposition 3.145 gives the result. □

Lemma 3.103.
The simple roots are linearly independent.

Proof. In the definition of a simple root, we need an order notion on Δ which is then seen as a subset of $\mathfrak{h}_{\mathbb{R}}$. But the roots are real thereon. Then the right notion of "independence" for the simple root is the independence with respect to *real* combinations.

If one has a combination $c^i\alpha_i = 0$ (sum over i) with at least one non zero among the c^i's by putting the negative c^i's at right, one can write

$$a^i\alpha_i = b^j\alpha_j$$

with $a^i, b^j \geqslant 0$. Let us consider $\gamma = a^i\alpha_i$ and h_γ. For every $h \in \mathfrak{h}$, we have

$$B(h, h\gamma) = \gamma(h) = a^i\alpha_i(h_\gamma).$$

but $h_\gamma = a^j h_\alpha$, then

$$B(h_\gamma, h_\gamma) = a^i a^j \alpha_i(h_{\alpha_j}) = a^i a^j B(h_{\alpha_i}, h_{\alpha_i}). \tag{3.157}$$

Since the α_i are all simple roots, the right hand side is negative, but proposition 3.145 makes the left hand side positive. Thus $\gamma = 0$. □

Theorem 3.104.
If $\{\alpha_1, \ldots, \alpha_r\}$ is the set of all the simple roots, then $\dim \mathfrak{h}_{\mathbb{R}} = r$ and every $\beta \in \Phi$ can be decomposed as

$$\beta = \sum_{i=1}^{r} n_i\alpha_i$$

where the n_i are integers either all positive either all negative.

Proof. Let β be a non simple positive root. Then it can be decomposed as $\beta = \gamma + \delta$ with $\gamma, \delta > 0$.We can also separately decompose γ and δ and continue so until we are left with simple roots. We have to see why the process stops. Since there are only a finite number of positive root, if the process does not stop, then the decomposition of (at least) one of the positive roots γ contains γ itself. So we have a situation $\gamma = \gamma + \alpha$ for a certain positive α. This contradict the notion of order.

In particular $\mathrm{Span}_{\mathbb{N}}\{\alpha_i\} = \{\text{positive roots}\}$. Thus it is clear that

$$\mathrm{Span}_{\mathbb{R}}\{\alpha_i\} = \Phi.$$

□

3.8.8 Weyl group

References about Weyl group: [6]. See also [19], page 530.

If α is a root of $\mathfrak{g}$ we define the **symmetry** of α as

$$\begin{aligned} s_\alpha \colon \mathfrak{h}^* &\to \mathfrak{h}^* \\ \beta &\mapsto \beta - \beta(H_\alpha)\alpha \end{aligned} \tag{3.158}$$

where $H_\alpha \in \mathfrak{h}$ is the inverse root of α. Since $\alpha(H_\alpha) = 2$ we have $s_\alpha(\alpha) = -\alpha$. The group generated by the symmetries and the identity is the **Weyl group**.

From what is said around equation (3.141) and the definition $(\alpha, \beta) = \alpha(t_\beta)$, we have

$$s_\alpha(\beta) = \beta - \frac{2(\alpha, \beta)}{(\alpha, \alpha)}\alpha. \tag{3.159}$$

We know from proposition 3.99 that $s_\alpha(\beta)$ is a root while there are only finitely many roots; thus the Weyl group is finite since there are only a finite number of maps from a finite set to itself.

The symmetries associated to roots are involutive:

$$s_\alpha^2 = \mathrm{id}\,. \tag{3.160}$$

Indeed

$$\begin{aligned} s_\alpha^2(\beta) &= s_\alpha\big(\beta - \beta(H_\alpha)\alpha\big) \\ &= \beta - \beta(H_\alpha) - \big(\beta - \beta(H_\alpha)\alpha\big)(H_\alpha)\alpha \\ &= \beta \end{aligned} \tag{3.161}$$

if we take into account $\alpha(H_\alpha) = 2$.

Relative to the symmetry s_{α_i} we have the symmetry s_i on $\mathfrak{h}$ defined by

$$s_i(h) = h - \alpha_i(h)H_i \tag{3.162}$$

where $h \in \mathfrak{h}$ and H_i is the inverse root of α_i.

Remark 3.105.
The simple roots α_i **are not** *orthogonal.*

Let Δ be a reduced abstract root system on a real finite dimensional vector space V. The group W generated by the $s_\alpha : \alpha \in \Delta$ is the **Weyl group**.

Proposition 3.106.
The elements s_{α_i} are isometries of $\mathfrak{h}^*$, i.e.

$$\big(s_{\alpha_i}(\alpha), s_{\alpha_i}(\beta)\big) = (\alpha, \beta). \tag{3.163}$$

Proof. For the sake of shortness, let us write

$$n_{i,\alpha} = \frac{2(\alpha_i, \alpha)}{(\alpha_i, \alpha_i)}. \tag{3.164}$$

We have $t_{s_{\alpha_i}(\alpha)} = t_\alpha - n_{i,\alpha} t_{\alpha_i}$. Thus

$$B\big(t_{s_{\alpha_i}(\alpha)}, t_{s_{\alpha_i}(\beta)}\big) = B(t_\alpha - n_{i,\alpha}t_{\alpha_i}, t_\beta - n_{i,\beta}t_{\alpha_i}) \tag{3.165}$$

distributing and taking into account the fact all the relations like $B(t_\alpha, t_{\alpha_i}) = (\alpha, \alpha_i)$, the right hand side reduces to $B(t_\alpha, t_\beta) = (\alpha, \beta)$. □

When Φ is the root system, one can chose many different notions of positivity; each of them bring to different simple systems. It turns out that the action of the Weyl group on a simple system produces the simple system of an other choice of positivity on Φ.

Lemma 3.107.
If $s_{\alpha_i}\alpha = s_{\alpha_i}\beta$, then $\alpha = \beta$.

Proof. The hypothesis $s_{\alpha_j}(\alpha - \beta) = 0$ provides

$$0 = \alpha - \beta - \frac{2(\alpha - \beta, \alpha_j)}{(\alpha_j, \alpha_j)}\alpha_j \tag{3.166}$$

so that $\alpha = \beta + z\alpha_j$ for some $z \in \mathbb{C}$. Thus we have

$$s_{\alpha_j}(\alpha) = s_{\alpha_j}(\beta) + z s_{\alpha_j}(\alpha_j) = s_{\alpha_j}(\alpha) - z\alpha_j. \tag{3.167}$$

Thus $z = 0$ and $\alpha = \beta$. □

Proposition 3.108.
Let α_i a simple root. The set $\Phi^+ \backslash \{\alpha_i\}$ is stable under s_{α_i}, i.e.

$$s_{\alpha_i}\big(\Phi^+ \backslash \{\alpha_i\}\big) = \Phi^+ \backslash \{\alpha_i\}. \tag{3.168}$$

Proof. Let $\lambda \in \Phi^+$ be a positive root. By theorem 3.101 we have

$$\lambda = \sum_j a_j \alpha_j \tag{3.169}$$

with $a_j \geqslant 0$. Since $\lambda \neq \alpha_i$ we have $a_j > 0$ for some $j \neq i$. Indeed the only multiple of α_i to be a root are 0 and $\pm\alpha_i$. Since $\lambda \in \Phi^+$ and $\lambda \neq \alpha_i$, none of these three solutions are taken into consideration.

Let's apply s_{α_i} on both sides of (3.169):

$$\begin{aligned} s_{\alpha_i}(\lambda) &= s_{\alpha_i}\big(\sum_j a_j \alpha_j\big) \\ &= \sum_{j \neq i} a_j s_{\alpha_i}(\alpha_j) + a_i \underbrace{s_{\alpha_i}(\alpha_i)}_{-\alpha_i} \\ &= \sum_{j\neq i} a_j \alpha_j - \sum_{j \neq i} a_j \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}\alpha_i - a_i\alpha_i \end{aligned} \tag{3.170}$$

Since a root is either positive or negative, the coefficients are either *all* positive or *all* negative. Since all the coefficients (apart for the one of α_i) are the same as the ones of λ, the root (3.170) is positive.

We still have to prove that $s_{\alpha_i}(\lambda) \neq \alpha_i$. Indeed if $s_{\alpha_i}(\lambda) = \alpha_i$ we have

$$\lambda = s_{\alpha_i} s_{\alpha_i}(\lambda) = s_{\alpha_i}(\alpha_i) = -\alpha_i, \tag{3.171}$$

which contradicts the positivity of λ.

Up to now we proved that $s_{\alpha_i}\big(\Phi^+ \backslash \{\alpha_i\}\big) \subset \Phi^+ \backslash \{\alpha_i\}$. If $\lambda \in \Phi^+ \backslash \{\alpha_i\}$, then

$$\sigma = s_{\alpha_i}(\lambda) \in s_{\alpha_i}\big(\Phi^+ \backslash \{\alpha_i\}\big) \subset \Phi^+ \backslash \{\alpha_i\} \tag{3.172}$$

and $s_{\alpha_i}(\sigma) = \lambda$, so that λ is the image by s_{α_i} of $\sigma \in \Phi^+ \backslash \{\alpha_i\}$. □

Theorem 3.109.
The map $s_{\alpha_j} : \Phi^+ \backslash \{\alpha_j\} \to \Phi^+ \backslash \{\alpha_j\}$ is bijective.

Proof. Surjectivity is proposition 3.108 while injectivity is lemma 3.107. □

Lemma 3.110 ([19], page 533)**.**
We consider the half sum of the positive roots:

$$\delta = \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha. \tag{3.173}$$

We have

(i) If α_j is a simple root, $s_{\alpha_j}\delta = \delta - \alpha_j$.

(ii) If α_j is a simple root, $(\delta, \alpha_j) = \frac{1}{2}(\alpha_j, \alpha_j)$.

Proof. We compute $s_{\alpha_j}\delta$ dividing the sum into two parts:

$$s_{\alpha_j}\delta = \frac{1}{2} \sum_{\substack{\alpha \in \Phi^+ \\ \alpha \neq \alpha_j}} s_{\alpha_j}(\alpha) + \frac{1}{2} s_{\alpha_j}(\alpha_j) \tag{3.174a}$$

$$= \frac{1}{2} \sum_{\substack{\alpha \in \Phi^+ \\ \alpha \neq \alpha_j}} \alpha - \frac{1}{2}\alpha_j. \tag{3.174b}$$

The second inequality is from the fact that s_{α_j} is bijective on $\Phi^+ \backslash \{\alpha_j\}$ by theorem 3.109. Adding a subtracting $\frac{\alpha_j}{2}$ we get

$$s_{\alpha_j}\delta = \frac{1}{2} \sum_{\alpha \in \Phi^+} \alpha - \frac{\alpha_j}{2} - \frac{\alpha_j}{2} = \delta - \alpha_j \tag{3.175}$$

Using the proposition 3.106, we have

$$(\delta, \alpha_j) = (s_{\alpha_j}\delta, s_{\alpha_j}\alpha_j) = (\delta - \alpha_j, -\alpha_j) = -(\delta, \alpha_j) + (\alpha_j, \alpha_j), \tag{3.176}$$

consequently, $2(\delta, \alpha_j) = (\alpha_j, \alpha_j)$ and the result follows. □

3.8.9 Abstract root system

The material about abstract root system mainly comes from [6].

Definition 3.111.
*An **abstract root system** in a finite dimensional vector space V endowed with an is a subset Φ of V such that*

- *Φ is finite and* $\operatorname{Span} \Phi = V$*,*
- *For every $\alpha \in \Phi$, there is a symmetry s_α of vector α leaving Φ stable.*
- *For every $\alpha, \beta \in \Phi$, the vector $s_\alpha(\beta) - \beta$ is an integer multiple of α.*

The abstract system is **reduced** when $\alpha \in \Phi$ implies $2\alpha \notin \Phi$. It is **irreducible** is Φ doesn't admits non trivial decomposition as $\Phi = \Phi' \cup \Phi''$ with $(\alpha, \beta) = 0$ for any $\alpha \in \Phi'$ and $\beta \in \Phi''$. We use the notation $\Phi := \Phi \cup \{0\}$.

The following is a consequence of all we did up to now.

Theorem 3.112.
The root system of a complex semisimple Lie algebra is a reduced abstract root system.

The **Weyl group** of Φ is the subgroup of $\mathrm{GL}(V)$ generated by the transformations s_α with $\alpha \in \Phi$.

3.8.9.1 Link with other definitions

The definition 3.111 is not the "usual" one (in [16], page 14 for example). We show now that we retrieve the usual features of an abstract.

Lemma 3.113.
An abstract root system admits a bilinear positive symmetric non degenerate form which is invariant under its Weyl group.

Proof. If $(.,.)_1$ is a bilinear positive non degenerate symmetric form on the vector space V, the form

$$(\alpha, \beta) = \sum_{w \in W} (w\alpha, w\beta)_1 \tag{3.177}$$

is invariant under the Weyl group. This construction is possible since the Weyl group is finite. □

Definition 3.114.
Let V be a vector space and $v \in V$ a non vanishing vector. A symmetry of vector v is an automorphism $s\colon V \to V$ such that

(i) $s(v) = -v$;

(ii) the set $H = \{w \in V \text{ st } \alpha(w) = w\}$ *is an hyperplane in* V.

A symmetry of vector v induces the decomposition $V = H \oplus \mathbb{R}v$. The symmetries are of order 2: $s^2 = \mathrm{id}$.

Lemma 3.115.
let v be a nonzero vector of V and A be a finite part of V such that $\mathrm{Span}(A) = V$. *Then there exists at most one symmetry of vector v leaving A invariant.*

Proof. Let s and s' be two such symmetries and consider $u = ss'$. We immediately have $u(A) = A$ and $u(v) = v$. Let us prove that u induce the identity on the quotient $V/\mathbb{R}v$. A general vector in V can be written (in a non unique way) under the form

$$h + h' + v \tag{3.178}$$

with $h \in H$ and $h' \in H'$. Let $h = h'_1 + \beta v$ be the decomposition of h in $H' \oplus \mathbb{R}v$ and $h' = h_1 + \gamma v$ be the decomposition of h' with respect to the direct sum $V = H \oplus \mathbb{R}v$. Then we have

$$\begin{aligned} ss'(h + h' + \alpha v) &= ss'\big((h'_1 + \beta v) + h' + \alpha v\big) && (3.179a)\\ &= s\big((h'_1 - \beta v) + h' + \alpha v\big) && (3.179b)\\ &= s(h - 2\beta v + h_1 + \gamma v + \alpha v) && (3.179c)\\ &= h + 2\beta v + h_1 - \gamma v + \alpha v && (3.179d)\\ &= h + h' + (\alpha - 2\gamma + 2\beta)v. && (3.179e) \end{aligned}$$

Thus at the level of the quotient, u leaves invariant $h + h'$.

It is not guaranteed that u is the identity, but the eigenvalues of u are 1. For each $x_i \in A$, there exists $n_i \in \mathbb{N}$ such that $u^{n_i} x_i = x$. If n is a common multiple of all the n_i (these are finitely many), we have $u^n(x) = x$ for every $x \in A$. Since A generates V, we have $u^n = \mathrm{id}$ and then u is diagonalizable.

We already mentioned the fact that the eigenvalues of u are 1. Since u is diagonalizable, it is the identity and $s = s'$. □

The invariant form give to V a structure of euclidian vector space for which the elements of the Weyl group are orthogonal matrices. Thus the symmetries read

$$s_\alpha(x) = x - 2\frac{(x, \alpha)}{(\alpha, \alpha)}\alpha. \tag{3.180}$$

This is the only transformation which makes $s_\alpha(\alpha) = -\alpha$ in the same time as being implemented by an orthogonal matrix. The symmetry s_α is nothing else than the orthogonal symmetry with respect to the hyperplane orthogonal to α.

The expression (3.180) has the consequence that

$$s_\alpha(\beta) - \beta = -\frac{(\beta, \alpha)}{(\alpha, \alpha)}\alpha. \tag{3.181}$$

By the definition of an abstract root system, the latter has to be an integer multiple of α, so

$$\frac{2(\beta,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}. \tag{3.182}$$

Definition 3.116.
Two abstract root systems Φ on V and Φ' on V' are ***isomorphic*** *is there exists an isomorphism of vector space $\psi\colon V \to V'$ such that $\psi(\Phi) = \Phi'$ and*

$$2\frac{(\alpha,\beta)}{(\alpha,\alpha)} = 2\frac{\big(\psi(\alpha),\psi(\beta)\big)}{\big(\psi(\alpha),\psi(\alpha)\big)} \tag{3.183}$$

for every $\alpha,\beta \in \Phi$.

3.8.9.2 Basis of abstract root system

The part about basis of abstract root system comes from [18].

Definition 3.117.
Let Φ be an abstract root system. A part $S \subset \Phi$ is a ***basis*** *of Φ if*

(i) S is a basis of V as vector space;

(ii) every $\beta \in \Phi$ can be written under the form

$$\beta = \sum_{\alpha \in S} m_\alpha \alpha \tag{3.184}$$

where m_α are all integers of the same sign.

The set Δ of simple roots of the root system of a complex semisimple Lie algebra is a basis.

We are going to build a basis of an abstract root system. Let $h \in V^*$ be such that $\alpha(h) \neq 0$ for every $\alpha \in \Phi$ and define

$$\Phi_h^+ = \{\alpha \in \Phi \text{ st } \alpha(h) > 0\}. \tag{3.185}$$

We have $\Phi = \Phi_h^+ \cup -\Phi_h^+$. We say that an element $\alpha \in \Phi_h^+$ is **decomposable** if there exist $\beta,\gamma \in \Phi_h^+$ such that $\alpha = \beta + \gamma$. We write S_h the set of undecomposable elements in Φ_h^+.

Lemma 3.118.
Any element in Φ_h^+ is a linear combination with positive coefficients of elements of S_h.

Problem and misunderstanding 10.
It seems to me that Serre's book[18] has a misprint here. At page V-11 he writes :

> *Tout élément de R_t^+ est combinaison linéaire, à coefficients entiers $\geqslant 0$ des éléments de S.*

Shouldn't he have written S_t.

Proof. Let I be the set of $\alpha \in \Phi_h^+$ that cannot be written under such a decomposition. We choose $\alpha \in I$ such that $\alpha(h)$ is minimal. If α is undecomposable, then $\alpha \in S_h$ and the condition $\alpha \in I$ is contradicted. Thus α is decomposable. Let $\beta,\gamma \in \Phi_h^+$ be such that $\alpha = \beta + \gamma$. Since $\alpha(h)$ is minimal,

$$\begin{aligned} \beta(h) &\leqslant \alpha(h) \\ \gamma(h) &\leqslant \alpha(h). \end{aligned} \tag{3.186}$$

Thus we have $\beta(h) = \alpha(h) - \gamma(h) < 0$ which contradicts $\beta \in \Phi^+$. We conclude that I is empty. □

Lemma 3.119.
If $\alpha,\beta \in S_h$, then $(\alpha,\beta) \leqslant 0$.

Proof. If $(\alpha,\beta) \geqslant 0$, then proposition 3.125(v) shows that $\gamma = \alpha - \beta$ is a root. There are two possibilities: $\gamma \in \pm\Phi_h^+$. If $\gamma \in \Phi_h^+$, then $\alpha = \gamma + \beta$ is decomposable; contradiction. If $\gamma \in -\Phi_h^+$, then $\beta = \alpha - \gamma$ is decomposable; contradiction. □

Lemma 3.120 (Lemme 4 page V-12)**.**
Let $h \in V^*$ and $A \subset V$ be a subset satisfying

(i) $\alpha(h) > 0$ for every $\alpha \in A$;

(ii) $(\alpha, \beta) \leqslant 0$ *for every* $\alpha, \beta \in A$.

Then the elements in A *are linearly independent.*

Proof. Let us consider a vanishing linear combination of elements in A:

$$\sum_{\alpha \in A} m_\alpha \alpha = 0. \tag{3.187}$$

We can sort the terms following that m_α is positive or negative and cut the sum in two parts:

$$\sum_{\beta \in A_1} y_\beta \beta = \sum_{\gamma \in A_2} z_\gamma \gamma \tag{3.188}$$

with $y_\beta, z_\gamma \geqslant 0$ and where A_1 and A_2 are disjoint subsets of A. Let us consider $\lambda = \sum_{\beta \in A_1} y_\beta \beta$ and compute

$$(\lambda, \lambda) = \sum_{\substack{\beta \in A_1 \\ \gamma \in A_2}} y_\beta z_\gamma (\beta, \gamma). \tag{3.189}$$

By hypothesis (β, γ) is lower than zero and by construction the product y_β, z_γ is positive. Thus the right hand side of equation (3.189) is negative. We conclude that $\lambda = 0$. Thus

$$0 = \lambda(h) = \sum_{\beta \in A_1} y_\beta \beta(h). \tag{3.190}$$

Since all the terms in the sum are larger than zero we have $y_\beta = 0$. In the same way we get $z_\gamma = 0$. The vanishing linear combination (3.187) is then trivial and the elements of A are linearly independent. □

Proposition 3.121.
The elements of S_h *form a basis of* Φ *in the sense of definition 3.117. Conversely, if* S *is a basis of* Φ *and if* $h \in V^*$ *is such that* $\alpha(h) > 0$ *for every* $\alpha \in S$*,we have* $S = S_h$.

Proof. The set S_h satisfies the conditions of lemma 3.120 since by definition $\alpha(h) > 0$ for every $\alpha \in S_h$ and by lemma 3.119 the inner products are all negative. Thus S_h is a free set. It is generating by lemma 3.118. Again by lemma 3.118, every element in Φ can be written as sum of elements of S_h with all coefficients of the same sign. Here we use the fact that v is positive if and only if $-v$ is negative and that every vector is either positive or negative.

For the second part, let S be a basis and $h \in V^*$ such that $\alpha(h) > 0$ for all $\alpha \in S$. Let

$$\Phi^+ = \{\sum_{\alpha \in S} m_\alpha \alpha \text{ with } m_\alpha \in \mathbb{N}\}. \tag{3.191}$$

We have $\Phi^+ \subset \Phi_h^+$ and $-\Phi^+ \subset -\Phi_h^+$. Since $\Phi = \Phi^+ \cup -\Phi^+$ we also have $\Phi^+ = \Phi_h^+$. Since elements of S are indecomposable in Φ^+, they are indecomposable in Φ_h^+ and we have $S \subset S_h$.

The sets S and S_h have the same number of elements because they both are basis of V, thus $S = S_h$. □

Lemma 3.122.
If h *and* h' *are elements of* V^* *related by* $\alpha(h) = (w\alpha)h'$*, then* $w(S_h) = S_{h'}$ *(if these space can be defined).*

Proof. Let $\alpha \in S_h$. The element $w(\alpha)$ belongs to $\Phi_{h'}^+$ because

$$w(\alpha)h' = \alpha(h) > 0 \tag{3.192}$$

because $\alpha \in \Phi_h^+$. We still have to check that $w(\alpha)$ is undecomposable in $\Phi_{h'}^+$. If $w(\alpha) = \beta + \gamma$ with $\beta, \gamma \in \Phi_{h'}^+$, we have $\alpha = w^{-1}\beta + w^{-1}\gamma$. From the link between h and h' we have

$$(w^{-1}\beta)(h) = (ww^{-1}\beta)h' = \beta(h') > 0. \tag{3.193}$$

Thus $w^{-1}\beta \in \Phi_h^+$ which is a contradiction because we supposed that α is undecomposable. □

Lemma 3.123.
If $\alpha, \beta \in \Phi$ *and if* $w \in W_S$*, then* $s_{w(\beta)} = w \circ s_\beta \circ w^{-1}$.

Proof. Using the fact that the symmetries are isometries of the inner product,

$$s_{w(\beta)}(\alpha) = \alpha - \frac{\big(w(\beta), \alpha\big)}{\big(w(\beta), w(\beta)\big)} w(\beta) = \alpha - \frac{(\beta), w^{-1}\alpha}{(\beta, \beta)} w\beta. \tag{3.194}$$

Applying that to $w(\alpha)$ instead of α and applying w^{-1}, we have

$$
\begin{aligned}
w^{-1}s_{w(\beta)}\big(w(\alpha)\big) &= w^{-1}\left(w\alpha - \frac{(\beta, w^{-1}w\alpha)}{(\beta,\beta)}w\beta\right) && (3.195a)\\
&= \alpha - \frac{(\beta,\alpha)}{(\beta,\beta)}w^{-1}w\beta && (3.195b)\\
&= s_\beta(\alpha). && (3.195c)
\end{aligned}
$$

□

The following theorem is from [18], page V-16.

Theorem 3.124.
Let W be the Weyl group of the abstract root system Φ. Let S a basis of Φ and W_S the subgroup of W generated by s_α with $\alpha \in S$. Then

(i) for every $h \in V^*$, there exists $w \in W_S$ such that $(w\alpha)(h) \geqslant 0$ for every $\alpha \in S$.

(ii) If S' is a basis of Φ, the there exists a $w \in W_S$ such that $w(S') = S$.

(iii) For every $\beta \in \Phi$ there exists $w \in W_S$ such that $w(\beta) \in S$.

(iv) The group W is generated by the symmetries s_α with $\alpha \in S$.

Proof. For item (i), consider $h \in V^*$ and $\delta = \frac{1}{2}\sum_{\gamma\in S}\gamma$. Let $w \in W_S$ be such that $w(\delta)h$ is the largest possible[19]. If $\alpha \in S$ we have

$$w(\delta)h \geqslant ws_\alpha(\delta)h = w(\delta)h - w(\alpha)h, \tag{3.196}$$

so that $w(\alpha)h \geqslant 0$ for every $\alpha \in S$. This proves our first assertion.

We pass to point (ii). Let $h' \in V^*$ be such that $\alpha'(h') > 0$ for every $\alpha' \in S'$. By the first item there exists $w \in W_S$ such that

$$(w\alpha)(h') \geqslant 0 \tag{3.197}$$

for every $\alpha \in S$. In fact we even have $w\alpha h' > 0$ for every $\alpha \in S$. Indeed $w\alpha$ can be decomposed as $\sum_{\alpha'\in S'} m_{\alpha'}\alpha'$ where all the $m_{\alpha'}$ have the same sign. In this case

$$(w\alpha)h' = \sum_{\alpha'} m_{\alpha'}\alpha'(h') \neq 0 \tag{3.198}$$

because each of $\alpha'(h')$ is strictly positive while all the terms of the sum have the same sign. This means, by the way, that $S' = S_{h'}$ following the proposition 3.121.

We define $h \in V^*$ by the relation

$$\alpha(h) = (w\alpha)(h'). \tag{3.199}$$

By what we said in equation (3.198) we have $\alpha(h) > 0$ for every $\alpha \in S$, so that we have $S = S_h$. Finally by lemma 3.122, $w(S_h) = S_{h'}$.

We prove now the point (iii). For $\gamma \in \Phi$ we consider the hyperplane

$$L_\gamma = \{h \in V^* \text{ st } \gamma(h) = 0\}. \tag{3.200}$$

Consider a particular $\beta \in \Phi$ the hyperplanes L_γ with $\gamma \neq \pm\beta$ do not coincide with L_β and there is only finitely many of them, so there exists a $h_0 \in L_\beta$ such that h_0 do not belong to any L_γ for any $\gamma \neq \pm\beta$.

In particular we have $\beta(h_0) = 0$ and $\gamma(h_0) \neq 0$ for every $\gamma \in \Phi$, $\gamma \neq \pm\beta$. If we choose ϵ small enough, there exists h near from h_0 such that

$$
\begin{cases}
\beta(h) = \epsilon > 0 & (3.201a)\\
|\gamma(h)| > \epsilon \qquad \text{if } \gamma \neq \pm\beta. & (3.201b)
\end{cases}
$$

Let S_h be the basis associated with this h. We have $\beta \in S_h$. Indeed first $\beta(h) = \epsilon > 0$ and if $\beta = \gamma + \rho$, we would have

$$\gamma(h) = \beta(h) - \rho(h) = \epsilon - \rho(h) < 0, \tag{3.202}$$

so that β is undecomposable in Φ_h^+. Now from point (ii) there exists $w \in W_S$ such that $w(S_h) = S$. In particular $w(\beta) \in S$.

[19] We can consider that w because W is finite.

We turn our attention to the item (iv). We are going to prove that $W = W_S$. Since W is generated by the symmetries s_β ($\beta \in \Phi$), it is sufficient to prove that W_S generates the symmetries s_β.

Let $\beta \in \Phi$ and consider the element $w \in W_S$ such that $\alpha = w(\beta) \in S$. From lemma 3.123 we have

$$s_\alpha = s_{w(\beta)} = w \circ s_\beta \circ w^{-1}, \tag{3.203}$$

so that

$$s_\beta = w^{-1} \circ s_\alpha \circ w \in W_S. \tag{3.204}$$

□

What this theorem says in the case of complex semisimple Lie algebras is that if $\{\alpha_1, \ldots, \alpha_l\}$ is the set of simple roots, the symmetries s_{α_i} generate the Weyl group. Now, since any root can be mapped on a simple one using the Weyl group, any root can be recovered from a simple one acting with the Weyl group that is generated by the simple ones.

Thus one can determine all the roots from the data of the simple ones by computing $s_{\alpha_i}\alpha_j$ and then acting again with the s_{α_i} on the results and again and again. This is the fundamental reason from which the root system can be recovered for the Cartan matrix.

3.8.9.3 Properties

The main properties of an abstract root system are given in the following proposition.

Proposition 3.125.
If Φ is an abstract root system in a vector space V, one has the following properties:

(i) *If $\alpha \in \Phi$ then $-\alpha \in \Phi$.*

(ii) *If $\alpha \in \Phi$, the multiples of α which could also be in Φ are either $\pm\alpha$, or $\pm\alpha$ and $\pm 2\alpha$ or $\pm\alpha$ and $\pm\frac{1}{2}\alpha$.*

(iii) *If $\alpha\beta \in \Phi$ then $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ can take the nonzero values ± 1, ± 2, ± 3 or ± 4. The case ± 4 can only arise if $\beta = \pm 2\alpha$.*

(iv) *If $\alpha, \beta \in \Phi$ are not proportional each other and if $|\alpha| \leqslant |\beta|$, then $\frac{2(\beta,\alpha)}{(\beta,\beta)}$ equals 0 or ± 1.*

(v) *If $\alpha, \beta \in \Phi$ and $(\alpha, \beta) > 0$, then $\alpha - \beta \in \Phi$ and if $(\alpha, \beta) < 0$, the $\alpha + \beta \in \Phi$.*

(vi) *If $\alpha, \beta \in \Phi$ and neither $\alpha + \beta$ neither $\alpha - \beta$ belongs to Φ, then $(\alpha, \beta) = 0$.*

(vii) *If $\alpha \in \Phi$ and $\beta \in \Phi$, the $n \in \mathbb{Z}$ such that $\beta + n\alpha \in \Phi$ fulfils $-p \leqslant n \leqslant q$ for certain $p, q \geqslant 0$. Moreover there are no gap,*

$$p - q = \frac{2(\alpha, \beta)}{(\alpha, \alpha)},$$

and there are at most four roots in the set $\{\beta + n\alpha\}_{-p \leqslant n \leqslant q}$.

(viii) *If Φ is reduced,*

  (a) *If $\alpha \in \Phi$, the only multiples of α to lies in Φ are $\pm\alpha$,*

  (b) *If $\alpha \in \Phi$ and $\beta \in \Phi$, then $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ can be equal to 0, ± 1, ± 2 or ± 3.*

The proof will not use the fact that Φ spans V.

Proof. (i) $s_\alpha \alpha = -\alpha$.

(ii) If $\beta = c\alpha$ with $|c| < 1$, then

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = 2c$$

must belongs to $\mathbb{Z}$, then $c = 0, \pm\frac{1}{2}$. If $|c| > 1$, we use exactly the same with $\alpha = \frac{1}{c}\beta$, so that $\frac{1}{c} = 0; \pm\frac{1}{2}$. Now if 2α is a root, it is clear that $\frac{1}{2}\alpha$ can't be.

If Φ is reduced, the fact that $\frac{1}{2}\alpha \in \Phi$ implies that $\alpha \notin \Phi$, so that $\pm\frac{1}{2}\alpha$ is excluded if $\alpha \in \Phi$, under the same assumption, 2α is also excluded. This proves (viii)a.

(iii) The Schwartz inequality $|(\alpha, \beta)| \leqslant |\alpha||\beta|$ gives

$$\left| \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \frac{2(\beta, \alpha)}{(\beta, \beta)} \right| \leqslant 4.$$

The equality only holds for $\beta = c\alpha$. In this case, we just saw that $\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = 2c$ with $c = 2$ at most. If the equality is strict, then $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ and $\frac{2(\beta,\alpha)}{(\beta,\beta)}$ are two integers whose product is $\leqslant 3$. The possibilities are 0, ± 1, ± 2, ± 3.

(iv) If $|\alpha| \leqslant |\beta|$, then the following integer inequality holds:

$$\left|\frac{2(\alpha,\beta)}{(\alpha,\alpha)}\right| \leqslant \left|\frac{2(\beta,\alpha)}{(\beta,\beta)}\right|.$$

Since the product of the two is $\leqslant 3$, the smallest is 0 or 1.

(v) If $\beta = c\alpha$, then $c = \pm\frac{1}{2}, \pm 2, \pm 1$. All the cases are easy. If $(\alpha,\beta) > 0$, then $c > 0$ and $\alpha - \beta = \alpha - \frac{1}{2}\alpha = \frac{1}{2}\alpha$ or $\alpha - \beta = \alpha - 2\alpha = -\alpha$.

Then we can suppose that α and β are not proportional each other. We consider $\alpha, \beta \in \Phi$ and $(\alpha,\beta) > 0$ (the other case is proved in much the same way). We just saw in (iv) that $\frac{2(\beta,\alpha)}{(\beta,\beta)}$ could be equals to 0 or ± 1, then the fact that $(\alpha,\beta) > 0$ imposes $\frac{2(\beta,\alpha)}{(\beta,\beta)} = 1$, so that $s_\beta(\alpha) = \alpha - \beta$.

If $|\beta| \leqslant |\alpha|$, we use

$$s_\alpha(\beta) = \beta - \frac{2(\beta,\alpha)}{(\beta,\beta)}\alpha = \beta - \alpha, \tag{3.205}$$

(vi) is an immediate consequence of the previous point.

(vii) Let $-p$ and q be the smallest and the largest values of n such that $\beta + n\alpha \in \Phi$. They exist because Φ is a finite set. Suppose that there is a gap between r and s ($r < s - 1$), i.e. $\beta + r\alpha \in \Phi$, $\beta + s\alpha \in \Phi$, but $\beta + (r+1)\alpha, \beta + (s-1)\alpha \notin \Phi$.

By the point (v), $(\beta + r\alpha, \alpha) \geqslant 0$ and $(\beta + s\alpha, \alpha) \leqslant 0$. Making the difference between these two inequalities,

$$(r - s)|\alpha|^2 \geqslant 0,$$

then $r \geqslant s$, which contradict the definition of r and s. So there is no gap. Now let us compute

$$\begin{aligned} s_\alpha(\beta + n\alpha) &= \beta + n\alpha - \frac{2(\alpha, \beta + n\alpha)}{(\alpha,\alpha)}\alpha \\ &= \beta + n\alpha - \left(\frac{2(\alpha,\beta)}{(\alpha,\alpha)} + 2n\right)\alpha \\ &= \beta - n\alpha - \frac{2(\alpha,\beta)}{(\alpha,\alpha)}\alpha \in \Phi. \end{aligned} \tag{3.206}$$

Then for any n in $-p \leqslant n \leqslant q$,

$$-q \leqslant n + \frac{2(\alpha,\beta)}{(\alpha,\alpha)} \leqslant p.$$

With $n = q$, the second inequality gives $\frac{2(\alpha,\beta)}{(\alpha,\alpha)} \leqslant p - q$ while the first one with $n = -p$ gives $p - q \leqslant \frac{2(\alpha,\beta)}{(\alpha,\alpha)}$.

The last point is to check the length of the string of root. We can suppose $q = 0$ (i.e to look the string of $\beta - q\alpha$ instead of the one of α; of course this is the same), then the length is $p + 1$ and

$$p = \frac{2(\alpha,\beta)}{(\alpha,\alpha)}.$$

If α and β are not proportional, the point (iii) makes it equals at most to 3. If they are proportional, then the possibilities are $\alpha = \pm\beta, \pm\frac{1}{2}\beta, \pm 2\beta$. The string $\beta + n\alpha$ with $\alpha = \beta$ is at most $\{\beta, 2\beta\}$, if $\alpha = \frac{1}{2}\beta$, this is just $\{\beta\}$ and if $\alpha = 2\beta$, this is $\{\beta, -\beta\}$.

The proof is complete. □

3.8.10 Abstract Cartan matrix

The following proposition summarize the properties of the of the Cartan matrix.

Definition 3.126.
A matrix $(A_{ij})_{1 \leqslant i,j \leqslant l}$ satisfying the following conditions is an ***abstract Cartan matrix***

(i) $A_{ij} \in \mathbb{Z}$,

(ii) $A_{ii} = 2$,

(iii) $A_{ij} \leqslant 0$ *if* $i \neq j$,

(iv) $A_{ij} = 0$ *if and only if* $A_{ji} = 0$,

(v) *there exists a diagonal matrix D with positive coefficients such that DAD^{-1} is symmetric and positive defined.*

The classification of abstract Cartan matrix will be performed in subsection 3.8.11. The data of an abstract Cartan matrix defines an abstract root system. For a proof, see [20].

Proposition 3.127.
The Cartan matrix of a semisimple complex Lie algebra is an abstract Cartan matrix.

Proof. The first two points are already done. For the point (iii), note that the sign of (α, β) is not sure when α is any root. However here we are speaking of simple roots. Let us consider the root

$$\lambda = \alpha_i - \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}\alpha_i \tag{3.207}$$

Since it is a root, proposition 3.101 says that the coefficients in the decomposition in simple roots have to be all integer and of the same sign. Thus the combination $(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i)$ has to be negative.

The point (v) is also non trivial. Consider the diagonal matrix $D = \operatorname{diag}\big((\alpha_i, \alpha_i)\big)_{i=1,\ldots,l}$. We have

$$(DAD^{-1})_{ij} = \sum_{kl} D_{ik} A_{kl} (D^{-1})_{lj} \tag{3.208a}$$

$$= \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)^{1/2} (\alpha_j, \alpha_j)^{1/2}}. \tag{3.208b}$$

This is a symmetric matrix. In order to proof that this is positive defined, we are going to provide a matrix B such that $DAD^{-1} = BB^t$. Let $\{\lambda_i\}$ be an orthonormal basis of $\mathfrak{h}^*$ and consider the matrix b given by the decomposition of the simple roots in this basis:

$$\alpha_i = \sum_j b_{ij} \lambda_j. \tag{3.209}$$

In particular we have $(\alpha_i, \alpha_j) = \sum_k b_{ik} b_{jk}$. Then we consider the matrix

$$B_{ij} = \frac{b_{ij}}{(\alpha_i, \alpha_i)^{1/2}} \tag{3.210}$$

which is non degenerate since the α_i are simple and thus a re linearly independent. Small computation shows that

$$(BB^t)_{ij} = \sum_k \frac{b_{ik}}{(\alpha_i, \alpha_i)^{1/2}} \frac{b_{jk}}{(\alpha_j, \alpha_j)^{1/2}} \tag{3.211a}$$

$$= \frac{(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)^{1/2} (\alpha_j, \alpha_j)^{1/2}} \tag{3.211b}$$

$$= (DAD^{-1})_{ij}. \tag{3.211c}$$

But BB^t is positive defined, then DAD^{-1} is. □

3.8.11 Dynkin diagrams

The sources for Dynkin diagrams is [9, 16].

We are going to associate to each abstract Cartan matrix, a diagram that will uniquely correspond to an abstract root system. In other words what we are going to do is to classify the matrix satisfying the conditions of definition 3.126.

If A is an abstract Cartan matrix we build the **Dynkin diagram** of A with the following rules.

(i) We put l vertices (one for each root)

(ii) The vertex i and j are joined with $A_{ij}A_{ji}$ lines.

A step by step construction is available in [16].

In the following we are considering an abstract Cartan matrix A and its associated abstract root system $\{\alpha_i\}$.

Lemma 3.128.
A abstract Cartan matrix with its abstract root system and its Dynkin diagram have the following properties.

(i) *If one remove the ith line an column of an abstract Cartan matrix, one still has an abstract Cartan matrix.*

(ii) *Two vertices are linked by* at most *three lines.*

(iii) *Each Dynkin diagram has more vertices than linked pairs.*

(iv) *A Dynkin diagram has no loop.*

(v) *A vertex in a Dynkin diagram has at most three lines attached (including multiplicities). Note: this is a generalization of point (ii).*

(vi) *Two root linked by a simple edge have equal **weight**, that is $(\alpha_i, \alpha_i) = (\alpha_j, \alpha_j)$.*

(vii) *If the two roots α_i, α_i are connected by a simple edge, we can collapse them, removing the connecting edge and conserving all the other edges.*

Proof. For point (ii) we have

$$A_{ij}A_{ji} = 4\frac{(\alpha_i,\alpha_j)}{(\alpha_i,\alpha_i)}\frac{(\alpha_j,\alpha_i)}{(\alpha_j,\alpha_j)} < 4 \tag{3.212}$$

by Cauchy-Schwarz inequality. We insist on the fact that the inequality is strict since α_i and α_j are not collinear: they are simple roots.

For point (iii) consider the form

$$\gamma = \sum_{i=1}^{l} \alpha_i(\alpha_i,\alpha_i)^{1/2}. \tag{3.213}$$

Since the simple roots are linearly independent, this sum is nonzero and we have $0 < (\gamma, \gamma)$. We have

$$\begin{aligned} 0 < (\gamma,\gamma) &= \sum_{ij} \frac{(\alpha_i,\alpha_j)}{\sqrt{(\alpha_i,\alpha_i)(\alpha_j,\alpha_j)}} \\ &= 2\sum_{i<j} \frac{(\alpha_i,\alpha_j)}{\sqrt{(\alpha_i,\alpha_i)(\alpha_j,\alpha_j)}} + \text{number of nodes} \\ &= -\sum_{i<j}(A_{ij}A_{ji})^{1/2} + \text{number of nodes}. \end{aligned} \tag{3.214}$$

Since for each linked pair (i, j) we have a term $A_{ij}A_{ji} \geqslant 1$, the positivity of the sum shows that

$$\text{number of nodes} > \sum_{ij} A_{ij}A_{ji} \geqslant \text{number of pairs}. \tag{3.215}$$

For item (iv), suppose that a loop is given by the roots $\alpha_1, \ldots, \alpha_n$. Since any sub-Dynkin diagram is a Dynkin diagram (from point (i)), we can consider only the loop. This is a diagram with n vertices and n pairs, which contradicts point (iii).

We pass to item (v). Let α_0 be a root linked to n simple lines, m double lines and p triple lines. For notational convenience, we write $v_i = \alpha_i/(\alpha_i, \alpha_i)$, $\{v_i\}_{1\leqslant i\leqslant n}$ is the set of "simply" linked roots to α_0, $\{v_i'\}_{1\leqslant i\leqslant m}$ the set of "doubly" linked and $\{v_i''\}_{1\leqslant i\leqslant p}$ the set of "triply" ones. Consider the vector

$$\gamma = v_0 + \sum_{i=1}^{n} f_i v_i + \sum_{i=1}^{m} g_i v_i' + \sum_{i=1}^{p} h_i v_i'' \tag{3.216}$$

where f_i, g_i and h_i are constant to be determined. In order to compute the norm of γ, notice that since there are no loops, no lines join v_i, v_i' and v_i'' together, so we have $(v_i, v_j') = (v_i, v_j'') = (v_i', v_j'') = 0$ and from the number of lines, $(v_0, v_i) = -1/2$, $(v_0, v_i') = -1/\sqrt{2}$ and $(v_0, v_i'') = -\sqrt{3}/2$. Thus we have

$$(\gamma,\gamma) = 1 + \sum_{i=1}^{m}(f_i^2 - f_i) + \sum_{i=1}^{m}(g_i^2 - \sqrt{2}g_i) + \sum_{i=1}^{p}(h_i^2 - \sqrt{3}h_i). \tag{3.217}$$

The minimum is realised with $f_i = 1/2$, $g_i = \sqrt{2}/2$ and $h_i = \sqrt{3}/2$ and for these values we have

$$(\gamma,\gamma) = 1 - \frac{n + 2m + 3p}{4}. \tag{3.218}$$

Since the inner product has to be positive we must have $n + 2m + 3p < 4$, the is the number of lines issued from α_0 has to be lower or equal to 3.

In order to proof (vi), remark that if α_i and α_j are connected by a simple edge, then $A_{ij}A_{ji} = 1$, which is only possible with $A_{ij} = A_{ji} = -1$. In particular we have $2(\alpha_i, \alpha_j)/(\alpha_i, \alpha_i) = 2(\alpha_j, \alpha_i)/(\alpha_j, \alpha_j)$, which proves that $(\alpha_i, \alpha_i) = (\alpha_j, \alpha_j)$.

Proof of item (vii). Since the two roots have same weight, the item (vi) says that up to permutation the Cartan matrix has a block 2×2 looking like

$$\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. \tag{3.219}$$

The proposed move consist to replace that block with the 1×1 matrix (2). As an example,

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & -1 \\ 0 & -1 & 2 & 0 \\ 0 & -1 & 0 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}. \tag{3.220}$$

It is clear that the obtained matrix is still an abstract Cartan matrix. $\square$

From these properties we can deduce much constrains on the Dynkin diagrams. First, the only diagram containing a triple edge is

$$\alpha_1 \equiv\!\equiv \alpha_2 \tag{3.221}$$

Let pass to the diagrams with only simple and double edges. If there is a double, there cannot be a triple point: the following is impossible

$$\begin{tikzcd}[arrows=dash] & & & & \alpha_5 \\ \alpha_1 \arrow[r, Rightarrow, no head] & \alpha_2 \arrow[r] & \alpha_3 \arrow[r] & \alpha_4 \arrow[ur] \arrow[dr] & \\ & & & & \alpha_6 \end{tikzcd} \tag{3.222}$$

since collapsing the roots α_2, α_3 and α_4 should create a point with four edges. Thus a diagram with a double edge is only possible inside a straight chain. Let us study the diagram

$$\alpha_1 - \alpha_2 = \alpha_3 - \alpha_4 - \alpha_5 \tag{3.223}$$

Once again we denote $v_i = \alpha_i/|\alpha_i|$ and we consider the (non vanishing) vector

$$\gamma = v_1 + bv_2 + cv_3 + dv_4 + ev_5 \tag{3.224}$$

whose norm is given by

$$(\gamma, \gamma) = 1 + b^2 + c^2 + d^2 + e^2 - b - \sqrt{2}bc = cd = de. \tag{3.225}$$

Equating all the partial derivative to zero provides the point

$$b = 2 \quad c = \frac{3}{\sqrt{2}} \quad d = \sqrt{2} \quad e = \frac{1}{\sqrt{2}}. \tag{3.226}$$

One check that with these values $(\gamma, \gamma) = 0$ which is impossible. The diagram (3.223) is thus impossible. By the collapsing principle, all the diagrams of the form

$$\alpha_1 - \alpha_2 = \alpha_3 - \alpha_4 - \cdots - \alpha_l \tag{3.227}$$

are impossible. The only possible diagrams with double edge are thus

$$\alpha_1 - \alpha_2 = \alpha_3 - \alpha_4 \tag{3.228a}$$
$$\alpha_1 = \alpha_2 - \cdots - \alpha_l \tag{3.228b}$$
$$\alpha_1 - \alpha_2 - \cdots - \alpha_{l-1} = \alpha_l. \tag{3.228c}$$

The diagrams (3.228b) and (3.228c) are the same. They however do not completely determine the abstract Cartan matrix because the diagram (3.228c) induces an asymmetry between α_1 and α_2. The so written Dynkin diagram cannot distinguish between the matrices

$$\begin{pmatrix} 2 & -2 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & -1 & 0 \\ -2 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \tag{3.229}$$

Thus we split the diagram (3.228c) into

$$\alpha_1 \longrightarrow \alpha_2 \longrightarrow \cdots \longrightarrow \alpha_{l-1} \Longrightarrow \alpha_l. \tag{3.230a}$$

$$\alpha_1 \longrightarrow \alpha_2 \longrightarrow \cdots \longrightarrow \alpha_{l-1} \Longleftarrow \alpha_l. \tag{3.230b}$$

In which the arrow points to the biggest root. The first one means that $|\alpha_1| = \ldots = |\alpha_{l-1}| = 1$, $\alpha_l = 2$ while the second diagram means $|\alpha_1| = \ldots = |\alpha_{l-2}| = |\alpha_l| = 1$, $\alpha_{l-1} = 2$.

We'll have to come back on this point later in subsection 3.8.12. Notice that this is the only diagram on which that problem occurs.

We are left to study the diagrams with only single edge. The following diagram is the simplest possible one:

$$\alpha_1 \text{ --- } \alpha_2 \text{ --- } \cdots \text{ --- } \alpha_l. \tag{3.231}$$

We have to know under what conditions one can have a triple point. We already know that there can be only one triple point.

If a diagram has a triple point, then one of the branch is of length 1. Indeed if not we would have the following diagram:

$$\begin{array}{ccccccc} & & & & \alpha_2 & \text{---} & \alpha_5 \\ & & & \diagup & & & \\ \alpha_7 & \text{---} & \alpha_4 \text{ --- } \alpha_1 & & & & \\ & & & \diagdown & & & \\ & & & & \alpha_6 & \text{---} & \alpha_6 \end{array} \tag{3.232}$$

Looking at the vector $\gamma = 3v_1 + 2(v_2 + v_3 + v_4) + v_5 + v_6 + v_7$ provides $(\gamma, \gamma) = -3$ which is impossible. Thus the diagrams with a branch are straight chains with one unique triple point which has a branch of length one. The question is: where can happen that branch ? The diagram

$$\begin{array}{ccccccccccccc} \alpha_1 & \text{---} & \alpha_2 & \text{---} & \alpha_3 & \text{---} & \alpha_4 & \text{---} & \alpha_5 & \text{---} & \alpha_6 & \text{---} & \alpha_7 \\ & & & & & & | & & & & & & \\ & & & & & & \alpha_8 & & & & & & \end{array} \tag{3.233}$$

cannot happen since the corresponding vector $v_1 + 2v_2 + 3v_3 + 4v_4 + 3v_5 + 2v_6 + v_7 + 2v_8$ has norm zero. Thus on a triple point, one branch has one branch of length 1 and at least one other to be of length 1 or 2. It turns out that all the diagrams of the form

$$\begin{array}{ccccc} & & & & \alpha_{l-1} \\ & & & \diagup & \\ \alpha_1 \text{ --- } \cdots \text{ --- } & \alpha_{l-2} & & & \\ & & & \diagdown & \\ & & & & \alpha_l \end{array} \tag{3.234}$$

are possible. We are thus left with diagrams with a triple point with a branch of length 1 and a branch of length 2 :

$$\begin{array}{ccccccccccc} \alpha_1 & \text{---} & \alpha_2 & \text{---} & \alpha_3 & \text{---} & \alpha_5 & \text{---} & \cdots & \text{---} & \alpha_l \\ & & & & | & & & & & & \\ & & & & \alpha_4 & & & & & & \end{array} \tag{3.235}$$

The diagram with a branch of length 5

$$\begin{array}{ccccccccccccc} \alpha_1 & \text{---} & \alpha_2 & \text{---} & \alpha_3 & \text{---} & \alpha_5 & \text{---} & \alpha_6 & \text{---} & \alpha_7 & \text{---} & \alpha_8 \\ & & & & | & & & & & & & & \\ & & & & \alpha_4 & & & & & & & & \end{array} \tag{3.236}$$

does not exist. We achieve the proof of that fact using for example this code for sage:

```
----------------------------------------------------------------------
| Sage Version 4.7.1, Release Date: 2011-08-11                       |
```

```
| Type notebook() for the GUI, and license() for information.        |
----------------------------------------------------------------------
sage: a=[var('a'+str(i-1)) for i in range(1,11)]
sage: l=9
sage: a[1]=1
sage: squares = sum( [a[i]**2 for i in range(1,l+1)] )     # The sum goes to l
sage: lines = sum(  [a[i]*a[i+1] for i in range(1,l-1)  ]   )+a[3]*a[9]  # The sum goes up to l-2
sage: f=symbolic_expression(squares - lines)
sage: X = solve( [f.diff(a[i])==0 for i in range(2,l+1)],[ a[i] for i in range(2,l+1)  ]  )
sage: print X[0]
[a2 == 2, a3 == 3, a4 == (5/2), a5 == 2, a6 == (3/2), a7 == 1, a8 == (1/2), a9 == (3/2)]
sage: f(   *tuple(  [  X[0][i].rhs() for i in range(0,l-1)  ]  )   )
0
```

This proves that the vector $v_1 + 2v_2 + 3a_3 + \frac{5}{2}v_4 + 2v_5 + \frac{3}{2}v_6 + v_7 + \frac{1}{2}v_8 + \frac{3}{2}v_9$ has vanishing norm, which is impossible.

$\boxed{\textbf{Problem and misunderstanding}}$ **11.**
This code raises a deprecation warning that I'm not able to solve.

We are finally left with the diagrams with one triple point with one branch of length 1, one branch of length 2 and the third branch with length 1, 2, 3 or 4 :

$$\begin{tikzcd}[arrows=dash] \alpha_1 \arrow[r] & \alpha_2 \arrow[r] & \alpha_3 \arrow[r] \arrow[d] & \alpha_4 \\ & & \alpha_5 & \end{tikzcd} \tag{3.237a}$$

$$\begin{tikzcd}[arrows=dash] \alpha_1 \arrow[r] & \alpha_2 \arrow[r] & \alpha_3 \arrow[r] \arrow[d] & \alpha_4 \arrow[r] & \alpha_5 \\ & & \alpha_6 & & \end{tikzcd} \tag{3.237b}$$

$$\begin{tikzcd}[arrows=dash] \alpha_1 \arrow[r] & \alpha_2 \arrow[r] & \alpha_3 \arrow[r] \arrow[d] & \alpha_4 \arrow[r] & \alpha_5 \arrow[r] & \alpha_6 \\ & & \alpha_7 & & & \end{tikzcd} \tag{3.237c}$$

$$\begin{tikzcd}[arrows=dash] \alpha_1 \arrow[r] & \alpha_2 \arrow[r] & \alpha_3 \arrow[r] \arrow[d] & \alpha_4 \arrow[r] & \alpha_5 \arrow[r] & \alpha_6 \arrow[r] & \alpha_7 \\ & & \alpha_8 & & & & \end{tikzcd} \tag{3.237d}$$

In order to list all the possible complex semisimple Lie algebra, we have to check each of the left Dynkin diagrams if they give rise to an abstract Cartan matrix.

3.8.12 Example of reconstruction by hand

We turn now our attention on the difference between the two diagrams (3.230). The Cartan matrix of the diagram $\alpha_1 \text{———} \alpha_2 \Longrightarrow \alpha_3$ is given by

$$A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -2 \\ 0 & -1 & 2 \end{pmatrix}. \tag{3.238}$$

The diagonal matrix D of definition 3.126 is

$$D = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 2 \end{pmatrix} \tag{3.239}$$

and the length of the roots are $\|\alpha_1\| = \|\alpha_2\| = 1$ and $|\alpha_3| = 2$. Let us compute the angles between the roots. In order to compute (α_1, α_2) we look at A_{12}:

$$A_{12} = -1 = 2\frac{(\alpha_1, \alpha_2)}{(\alpha_1, \alpha_2)}, \tag{3.240}$$

and the same computation with A_{23} provides

$$(\alpha_1, \alpha_2) = -\frac{1}{2} \tag{3.241a}$$
$$(\alpha_2, \alpha_3) = -1 \tag{3.241b}$$

We compute all the roots using the theorem 3.124 which basically says that acting with the "simple" Weyl group W_S on the simple roots generates all the roots. On the first strike we have

$$\begin{aligned} s_1(\alpha_2) &= \alpha_2 + \alpha_1 & s_2(\alpha_1) &= \alpha_1 + \alpha_2 & s_3(\alpha_1) &= \alpha_1 \\ s_1(\alpha_3) &= \alpha\alpha_3 & s_2(\alpha_3) &= \alpha_3 + 2\alpha_2 & s_3(\alpha_2) &= \alpha_2 + \alpha_3. \end{aligned} \tag{3.242}$$

We discovered the roots $\alpha_2 + \alpha_1$, $\alpha_3 + 2\alpha_2$ and $\alpha_2 + \alpha_3$. Acting again on these roots by s_{α_1}, s_{α_2} and s_{α_3} the only new results are

$$\begin{aligned} s_1(\alpha_3 + \alpha_2) &= \alpha_1 + \alpha_2 + \alpha_3 \\ s_1(\alpha_3 + 2\alpha_2) &= 2\alpha_1 + 2\alpha_2 + \alpha_3. \end{aligned} \tag{3.243}$$

Acting again we find only one new root:

$$s_{\alpha_2}(\alpha_1 + \alpha_2 + \alpha_3) = \alpha_1 + 2\alpha_2 + \alpha_3. \tag{3.244}$$

We check that acting once again with the three simple roots on this last one does not brings new roots. Thus we have 9 positive roots. Adding the negative ones, we are left with 18 root spaces of dimension one. The Cartan algebra has dimension 3, so the algebra we are looking at has dimension 21.

Now take a look at the similar Dynkin diagram and its Cartan matrix:

$$\alpha_1 \;\text{———}\; \alpha_2 \Longleftarrow \alpha_3 \qquad\qquad A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -2 & 2 \end{pmatrix} \tag{3.245a}$$

The inner products are

$$\begin{aligned} |\alpha_1| = |\alpha_3| = 1, |\alpha_2| = 2 \\ (\alpha_1, \alpha_2) = -1/\sqrt{2}, (\alpha_2, \alpha_3) = -1 \end{aligned} \tag{3.246}$$

and the roots are

$$\alpha_1 \tag{3.247a}$$
$$\alpha_2 \tag{3.247b}$$
$$\alpha_3 \tag{3.247c}$$
$$\alpha_1 + \alpha_2 \tag{3.247d}$$
$$\alpha_2 + \alpha_3 \tag{3.247e}$$
$$\alpha_2 + 2\alpha_3 \tag{3.247f}$$
$$\alpha_1 + \alpha_2 + \alpha_3 \tag{3.247g}$$
$$\alpha_1 + \alpha_2 + 2\alpha_3 \tag{3.247h}$$
$$\alpha_1 + 2\alpha_2 + 2\alpha_3. \tag{3.247i}$$

We see that the inner products are already not the same. Notice that the roots are really different: it is not simply a renaming $\alpha_2 \leftrightarrow \alpha_3$.

Thus the two Dynkin diagrams (3.228c) are describing two different Lie algebras.

3.8.13 Reconstruction

The construction theorem is the following.

Theorem 3.129.
*Let R be an abstract root system in a complex vector space V^* and $\{\alpha_1, \ldots, \alpha_n\}$ be a basis of R. We denote by $H_i \in V$ the **inverse root** of α_i(i.e. $\alpha(H_\alpha) = 2$). We define the Cartan matrix*

$$A_{ij} = \alpha_j(H_i). \tag{3.248}$$

Let $\mathfrak{g}$ be the Lie algebra defined by the $3n$ generators X_i, Y_i, H_i and the relations

$$[H_i, H_j] = 0 \tag{3.249a}$$
$$[X_i, Y_j] = \delta_{ij} H_i \tag{3.249b}$$
$$[H_i, X_j] = A_{ij} X_j \tag{3.249c}$$
$$[H_i, Y_j] = -A_{ij} Y_j \tag{3.249d}$$

and, for $i \neq j$,

$$\mathrm{ad}(X_i)^{-A_{ij}+1}(X_j) = 0 \tag{3.250a}$$
$$\mathrm{ad}(Y_i)^{-A_{ij}+1}(Y_j) = 0. \tag{3.250b}$$

Then $\mathfrak{g}$ is a semisimple Lie algebra in which a Cartan subalgebra is generated by $H_1, \ldots, H_n$ and its root system is R.

A complete proof can be found in [18] at page VI-19. We are going to give some ideas.

We consider $\mathfrak{g}$, the Lie algebra generated by the elements H_i, X_i and Y_i. We denote by $\mathfrak{h}$ the abelian Lie algebra generated by the elements H_i.

Lemma 3.130.
The endomorphism $\mathrm{ad}(X_i)$ and $\mathrm{ad}(Y_i)$ are nilpotent.

Proof. Let V_i the subspace of $\mathfrak{g}$ of elements z such that $\mathrm{ad}(X_i)^k z = 0$ for some $k \in \mathbb{N}$. The space V_i is a Lie subalgebra of $\mathfrak{g}$ because

$$\mathrm{ad}(X_i)[z, z'] = -[z, \mathrm{ad}(X_i)z'] + [z', \mathrm{ad}(X_i)z]. \tag{3.251}$$

Acting with $\mathrm{ad}(X_i)^n$ we get terms of the form $\big[\mathrm{ad}(X_i)^k z, \mathrm{ad}(X_i)^l z'\big]$ with $k + l = n$. If n is large enough, all the terms vanish.

From the relation (3.250a) we see that $X_j \in V_i$ for every j. Since $[X_i, H_j]$ is proportional to X_i we also have $H_j \in V_i$ and then $Y_j \in V_i$ because $[X_i, Y_j] = \delta_{ij} H_i \in V_i$. Thus the Lie algebra V_i contains all the Chevalley generators and then $V_i = \mathfrak{g}$. □

For $\lambda \in \mathfrak{h}^*$ we define

$$\mathfrak{g}_\lambda = \{z \in \mathfrak{g} \text{ st } \mathrm{ad}(h)z = \lambda(h)z \forall h \in \mathfrak{h}\}. \tag{3.252}$$

Then one prove that $\dim \mathfrak{g}_{\alpha_i} = 1$ and $\dim \mathfrak{g}_{m\alpha_i} = 0$ if $m \neq \pm 1, 0$. This corresponds to the fact that we have a reduced root system, which is always the case in complex semisimple Lie algebras[20]. We denote by Φ the subset of $\lambda \in \mathfrak{h}^*$ such that $\mathfrak{g}_\lambda \neq 0$.

It turns out that we have the direct sum decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha. \tag{3.253}$$

One of the key ingredients in this building is the following lemma.

Lemma 3.131.
If λ and μ are related by an element of the Weyl group, then $\dim \mathfrak{g}_\lambda = \mathfrak{g}_\mu$.

Proof. Lemma 3.130 allows us to introduce the automorphism

$$\theta_i = e^{\mathrm{ad}(X_i)} e^{-\mathrm{ad}(Y_i)} e^{\mathrm{ad}(X_i)} \tag{3.254}$$

of $\mathfrak{g}$. We see that the restriction of θ_i to $\mathfrak{h}$ is the symmetry associated to α_i (see (3.162)). Indeed the first exponential reduces to

$$e^{\mathrm{ad}(X_i)} H_k = H_k - A_{ki} X_i \tag{3.255}$$

where $A_{ki} = \alpha_i(H_k)$. The second exponential gives

$$\begin{aligned} e^{\mathrm{ad}(-Y_i)}(H_k - A_{ki}X_i) &= H_k - A_{ki}X_i + (-A_{ki}Y_i - A_{ki}H_i) + \frac{1}{2}(2A_{ki}Y_i) \\ &= H_k - A_{ki}H_i - A_{ki}X_i. \end{aligned} \tag{3.256}$$

Notice the simplification of $A_{ki}Y_i$. The third exponential then provides the result (after some simplifications):

$$e^{\mathrm{ad}(X_i)}(H_k - A_{ki}H_i - A_{ki}X_i) = H_k - A_{ki}H_i = H_k - \alpha_i(H_k)H_i. \tag{3.257}$$

[20] However, at this point we have not proved yet that $\mathfrak{g}$ is semisimple and has that root system.

We proved that $\theta_i(H_k) = s_I(H_k)$. We deduce that $\theta_i e_\alpha \in \mathfrak{g}_{s_{\alpha_i}(\alpha)}$ whenever $e_\alpha \in \mathfrak{g}_\alpha$. Since θ_i is an automorphism of $\mathfrak{g}$ we have

$$[H_k, \theta_i e_\alpha] = \theta_i[\theta_i^{-1} H_k, e_\alpha]. \tag{3.258}$$

Since θ_i reduces to the involutive automorphism s_i on $\mathfrak{h}$ we have $\theta_i^{-1} H_k = \theta_i H_k = s_i(H_k)$. Then we have

$$[H_k, \theta_i e_\alpha] = \theta_i[s_i(H_k), e_\alpha] = \theta_i \alpha\big(s_i(H_k)\big) e_\alpha. \tag{3.259}$$

The eigenvalue of $\theta_i e_\alpha$ for $\operatorname{ad}(H_k)$ is thus $\alpha\big(s_i(H_k)\big)$. Using the definition and $A_{ki} = \alpha_i(H_k)$ we have

$$\begin{aligned}
\alpha\big(s_i(H_k)\big) &= \alpha(H_k) - \alpha_i(H_k)\alpha(H_i)\\
&= \big(\alpha - \alpha(H_i)\alpha_i\big) H_k\\
&= s_{\alpha_i}(\alpha) H_k.
\end{aligned} \tag{3.260}$$

At the end we got

$$[H_k, \theta_i e_\alpha] = s_{\alpha_i}(H_k)\theta_i e_\alpha \tag{3.261}$$

and then $\theta_i e_\alpha \in \mathfrak{g}_{s_{\alpha_i}(\alpha)}$. Thus the automorphism θ_i transforms $\mathfrak{g}_\lambda$ into $\mathfrak{g}_\mu$ when $\mu = s_i(\lambda)$ and

$$\dim \mathfrak{g}_\lambda = \dim \mathfrak{g}_{s_i(\lambda)}. \tag{3.262}$$

□

From here we prove that $\dim \mathfrak{g}_\alpha = 1$ for every root α[21].

Now if $\alpha + \beta = \gamma + \mu$, the elements $[E_\alpha, E_\beta]$ and $[E_\gamma, E_\mu]$ are proportional since they belong to the one-dimensional space $\mathfrak{g}_{\alpha+\beta}$.

Remark 3.132.
A linear map $\phi\colon \mathfrak{g} \to V$ from $\mathfrak{g}$ to a vector space V can be defined on the generators X_i, Y_i and H_i among with a formula giving $\phi([X,Y])$ in terms of $\phi(X)$ and $\phi(Y)$.

Problem and misunderstanding 12.
This remark could be made more precise. I'm thinking to the proposition **??** *giving the standard bialgebra structure on a Lie algebra.*

3.8.14 Cartan-Weyl basis

Let us study the eigenvalue equation

$$\operatorname{ad}(A)X = \rho X. \tag{3.263}$$

The number of solutions with $\rho = 0$ depends on the choice of $A \in \mathfrak{g}$.

Lemma 3.133.
If A is chosen in such a way that $\operatorname{ad}(A)X = 0$ has a maximal number of solutions, then the number of solutions is equal to the rank of $\mathfrak{g}$ and the eigenvalue $\alpha = 0$ is the only degenerated one in equation (3.263).

We suppose A to be chosen in order to fulfill the lemma. Thus we have linearly independent vectors H_i $(i = 1, \ldots l)$ such that

$$[A, H_i] = 0 \tag{3.264}$$

where l is the rank of $\mathfrak{g}$. Since $[A, A] = 0$, the vector A is a combination $A = \lambda^i H_i$. Since $\operatorname{ad}(A)$ is diagonalisable, one can find vectors E_α with

$$[A, E_\alpha] = \alpha E_\alpha, \tag{3.265}$$

and such that $\{H_i, E_\alpha\}$ is a basis of $\mathfrak{g}$. Using the fact that $\operatorname{ad}(A)$ is a derivation, we find

$$[A, [H_i, E_\alpha]] = \alpha[H_i, E_\alpha], \tag{3.266}$$

The eigenvalue $\alpha = 0$ being the only one to be degenerated, one concludes that $[H_i, E_\alpha]$ is a multiple of E_α:

$$[H_i, E_\alpha] = \alpha_i E_\alpha. \tag{3.267}$$

Replacing $A = \lambda^i H_i$, we have

$$\alpha E_\alpha = [\lambda^i H_i, E_\alpha] = \lambda^i \alpha_i E_\alpha, \tag{3.268}$$

[21][18] page VI-23. Be careful: this is not the statement of page VI-2.

thus $\alpha = \lambda^i\alpha_i$ (with a summation over $i = 1,\ldots,l$).

Before to go further, notice that the space spanned by $\{H_i\}_{i=1,\ldots,l}$ is a maximal abelian subalgebra of $\mathfrak{g}$, so that it is a Cartan subalgebra that we, naturally denote by $\mathfrak{h}^*$. Thus, what we are doing here is the usual root space construction. In order to stick the notations, let us associate the form $\sigma_\alpha \in \mathfrak{h}^*$ defined by $\sigma_\alpha(H_i) = \alpha_i$. In that case,

$$\sigma_\alpha(A) = \sigma_\alpha(\lambda^i H_i) = \lambda^i\alpha_i = \alpha \tag{3.269}$$

and we have

$$[A, E_\alpha] = \sigma_\alpha(A)E_\alpha. \tag{3.270}$$

On the other hand, we have $[H_i, E_\alpha] = \alpha_i E_\alpha = \sigma_\alpha(H_i)E_\alpha$, so that the eigenvalue α is identified to the root α, and we have $E_\alpha \in \mathfrak{g}_\alpha$.

Let us now express the vectors t_α in the basis of the H_i. The definition property is $B(t_\alpha, H_i) = \alpha(H_i) = \alpha_i$. If $t_\alpha = (t_\alpha)^i H_i$, we have

$$\alpha_i = B(t_\alpha, H_i) = B_{kl}(t_\alpha)^k \underbrace{(H_i)^l}_{=\delta^l_i} = B_{ki}(t_\alpha)^k. \tag{3.271}$$

If (B^{ij}) are the matrix elements of B^{-1}, we have

$$(l_\alpha)^l = \alpha_i B^{il} = \alpha^l \tag{3.272}$$

where α^l is defined by the second equality. Using proposition 3.80, we have

$$[E_\alpha, E_{-\alpha}] = B(E_\alpha, E_{-\alpha})\alpha^l H_l. \tag{3.273}$$

Thus one can renormalise E_α in such a way to have

$$\begin{aligned}
[H_i, H_j] &= 0,\\
[E_\alpha, E_{-\alpha}] &= \alpha^i H_i\\
[H_i, E_\alpha] &= \alpha_i E_\alpha = \alpha(H_i)E_\alpha\\
[E_\alpha, E_\beta] &= N_{\alpha\beta}E_{\alpha+\beta}
\end{aligned} \tag{3.274}$$

where the constant $N_{\alpha\beta}$ are still undetermined. A basis $\{H_i, E_\alpha\}$ of $\mathfrak{g}$ which fulfill these requirements is a basis of **Cartan-Weyl**.

3.8.15 Cartan matrix

We follow [21]. We denote by Π the system of simple roots of $\mathfrak{g}$. All the positive roots have the form

$$\sum_{\alpha\in\Pi} k_\alpha\alpha \tag{3.275}$$

with $k_\alpha \in \mathbb{N}$.

Theorem 3.134.
Let α and β be simple roots Thus

(i) $\alpha - \beta$ is not a simple root

(ii) we have

$$\frac{2(\alpha,\beta)}{(\alpha,\alpha)} = -p \tag{3.276}$$

where p is a strictly positive integer.

Partial proof. We are going to prove that $\frac{2(\alpha,\beta)}{(\alpha,\alpha)}$ is an integer. Let α and γ be non vanishing roots such that $\alpha + \gamma$ is not a root, and define

$$E'_{\gamma-j\alpha} = \operatorname{ad}(E_{-\alpha})^k E_\gamma \in \mathfrak{g}_{\gamma-k\alpha}. \tag{3.277}$$

Since there are a finite number of roots, there exists a minimal positive integer g such that $\operatorname{ad}(E_{-\alpha})^{g+1}E_\gamma = 0$. We define the constants μ_k (which depend on γ and α) by

$$[E_\alpha, E'_{\gamma-k\alpha}] = \mu_k E'_{\gamma-(k-1)\alpha}. \tag{3.278}$$

Using the definition of $E'_{\gamma-k\alpha}$ and Jacobi, one founds

$$\mu_k E'_{\gamma-(k-1)\alpha} = \big[E'_\alpha, [E_{-\alpha}, E'_{\gamma-(k-1)\alpha}]\big] = \alpha^i[H_i, E'_{\gamma-(k-1)\alpha}] + \mu_{k-1}E'_{\gamma-(k-a)\alpha}, \tag{3.279}$$

so that $\mu_k = \alpha^i\gamma_i - (k-1)\alpha^i\alpha_i + \mu_{k-1}$, and we have the induction formula

$$\mu_k = (\alpha, \gamma) - (k-1)(\alpha, \alpha) + \mu_{k-1} \tag{3.280}$$

for $k \geqslant 2$. If we define $\mu_0 = 0$, that relation is even true for $k = 1$. The sum for $k = 1$ to $k = j$ is easy to compute and we get

$$\mu_j = j(\alpha, \gamma) - \frac{j(j-1)}{2}(\alpha, \alpha). \tag{3.281}$$

Since $\mu_{g+1} = 0$, we have

$$(\alpha, \gamma) = g(\alpha, \alpha)/2, \tag{3.282}$$

and thus

$$\mu_j = \frac{j(g-j+1)(\alpha, \alpha)}{2}. \tag{3.283}$$

Let β be any root and look at the string $\beta + j\alpha$. There exists a maximal $j \geqslant 0$ for which $\beta + j\alpha$ is a root while $\beta + (j+1)\alpha$ is not a root. Now we consider $\gamma = \beta + j\alpha$ with that maximal j. Putting $\gamma = \alpha + j\beta$ in (3.282), one finds

$$(\alpha, \beta) = \frac{(g-2j)(\alpha, \alpha)}{2}, \tag{3.284}$$

and finally,

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} = g - 2j, \tag{3.285}$$

which is obviously an integer.

□

From the inner product on $\mathfrak{h}^*$, we deduce a notion of **angle**:

$$\cos(\theta_{\alpha,\beta}) = \frac{(\alpha, \beta)}{\sqrt{(\alpha, \alpha)(\beta, \beta)}}. \tag{3.286}$$

The **length** of the root α is the number $\sqrt{(\alpha, \alpha)}$.

Lemma 3.135.
If α and β are roots, then

$$\frac{2(\alpha, \beta)}{(\alpha, \alpha)} \in \mathbb{Z}, \tag{3.287}$$

and

$$\beta - \frac{2(\alpha, \beta)}{(\alpha, \alpha)} \tag{3.288}$$

is a root too.

If α and β are non vanishing, then the α-string which contains β contains at most 4 roots. Finally, the ratio

$$\frac{2(\alpha, \beta)}{(\alpha, \beta)} \tag{3.289}$$

takes only the values 0, ± 1, ± 2 or ± 3.

Let $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ be a system of simple roots. The **Cartan matrix** is the $l \times l$ matrix with entries

$$A_{ij} = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}. \tag{3.290}$$

Notice that, in the literacy, one find also the convention $A_{ij} = 2(\alpha_i, \alpha_j)/(\alpha_j, \alpha_j)$, as in [22], for example.

Lemma 3.136.
There exist positive rational numbers d_i such that

$$d_i A_{ij} = d_j A_{ji} \tag{3.291}$$

where A is the Cartan matrix.

Proof. The numbers are given by

$$d_i = \frac{(\alpha_i, \alpha_i)}{(\alpha_1, \alpha_1)}. \tag{3.292}$$

The relations (3.291) are easy to check using the definition (3.290). The fact that d_i is a strictly positive rational number comes from (3.276). □

Problem and misunderstanding 13.
I think that there is a property saying (something like) that A_{ij} is the larger integer k such that $\alpha_i + k\alpha_j$ is a root.

3.9 Other results

3.9.1 Abstract Cartan matrix

As before if we chose a basis $\{\varphi_1 \ldots \varphi_l\}$ of V, we can consider a lexicographic ordering on V. A root is **simple** when it is positive and can't be written as as sum of two positive roots. As in a non abstract case, abstract simple root also have the following property:

Proposition 3.137.
If $\dim V = l$, one has only l simple roots $\alpha_1, \ldots, \alpha_l$; they are linearly independent and if $\beta \in \Phi$ expands into $\beta = \sum c_j\alpha_j$, the c_j's all are integers and the non zero ones all have the same sign.

An ordering on V gives a notion of simple roots. The $l \times l$ matrix whose entries are

$$A_{ij} = \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}$$

is the **abstract Cartan matrix** of the abstract root system and the given ordering.

Theorem 3.138.
The main properties are

(i) $A_{ij} \in \mathbb{Z}$,

(ii) $A_{ii} = 2$,

(iii) if $i \neq j$, then $A_{ij} \leqslant 0$ and A_{ij} can only take the values $0, -1, -2$ *or* -3,

(iv) if $i \neq j$, $A_{ij}A_{ji} < 4$ (no sum),

(v) $A_{ij} = 0$ is and only if $A_{ji} = 0$,

(vi) $\det A$ *is integer and positive.*

Proof. The last point is the only non immediate one. The matrix A is the product of the diagonal matrix with entries $2/|\alpha_i|^2$ and the matrix whose entries are (α_i, α_j). The fact that the latter is positive definite is a general property of linear algebra. If $\{e_i\}$ is a basis of a vector space V, the matrix whose entry ij is given by (e_i, e_i) is positive definite. Indeed one can consider an orthonormal basis $\{f_i\}$ and a nondegenerate change of basis $e_i = B_{ik}f_k$. Then $(e_i, e_j) = (BB^t)_{ij}$. It is easy to see that for all $v \in V$, we have $(BB^t)_{ij}v^iv^j = \sum_k (v^iB_{ik})^2 > 0$.

The fact that the determinant is integer is simply the fact that this is a polynomial with integer variables. □

If we have an ordering on V we define Φ^+, the set of positive roots. From there, one can consider Π, the set of simple roots. Any element of Φ expands to a sum of elements of Π. Note that the knowledge of Π is sufficient to find Φ^+ back because $\alpha > 0$ implies $\alpha = \sum c_i\alpha_i$ with $c_i \geqslant 0$.

We can make this reasoning backward. Let us consider $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ be a basis of V such that any $\alpha \in \Phi$ expands as a sum of α_i with all coefficients of the same sign. Such a Π is a **simple system**. From such a Π, we can build a Φ^+ as the set of elements of the form $\alpha = \sum c_i\alpha_i$ with $c_i \geqslant 0$.

Proposition 3.139.
The so build Φ^+ is the set of positive roots for a certain ordering.

Proof. If we consider on V the lexicographic ordering with respect to the basis Π, a positive element $\alpha = \sum c_i\alpha_i$ has at least one positive coefficient among the c_i. If $\alpha \in \Phi$, we can say (by definition of Π) that in this case *all* the coefficients are positive, then the positive roots exactly form the set Φ^+. □

From now when we speak about a Φ^+, it will always be with respect to a simple system. The advantage is the fact that there are no more implicit ordering.

Lemma 3.140.
Let $\Pi = \{\alpha_1, \ldots, \alpha_l\}$ be a simple system and $\alpha \in \Phi^+$. Then

$$s_{\alpha_i} = \begin{cases} -\alpha_i & \text{if } \alpha = \alpha_i \\ > 0 & \text{if } \alpha \neq \alpha_i. \end{cases}$$

Proof. The first case is well know from a long time. For the second, compute

$$\begin{aligned} s_{\alpha_i}(\sum c_j\alpha_j) &= \sum_{j\neq i} c_j\alpha_j + c_i\alpha_i - 2c_i\alpha_i - \sum_{j\neq i}\frac{2c_j}{|\alpha_i|^2}(\alpha_j,\alpha_i)\alpha_i \\ &= \sum_{j\neq i} + \left(-\sum_{i\neq j}\frac{2c_j}{|\alpha_i|^2}(\alpha_j,\alpha_i)+c_i\right)\alpha_i. \end{aligned} \tag{3.293}$$

We see that between $\sum c_k\alpha_k$ and $s_{\alpha_i}(\sum c_k\alpha_k)$, there is just the coefficient of α_i which changes. Then if $\alpha \neq \alpha_i$, the positivity is conserved.

□

Proposition 3.141.
Let $\Pi = \{\alpha_1,\ldots,\alpha_l\}$ be a simple system. Then W is generate by the s_{α_i}'s. If $\alpha \in \Phi$, then there exists a $\alpha_i \in \Pi$ and $s \in W$ such that $s\alpha_j = \alpha$.

Proof. We denote by W' the group generate by the s_{α_i}'s; the purpose is to show that $W = W'$. We begin to show that if $\alpha > 0$, then $\alpha = s\alpha_j$ for certain $s \in W'$ and $\alpha_j \in \Pi$. For this, we write $\alpha = \sum c_j\alpha_j$ and we make an induction with respect to $\text{Level}(\alpha) = \sum c_j$. If $\text{Level}(\alpha) = 1$, then $\alpha - \alpha_j$ and $s = \text{id}$ works. Now we suppose that it works for Level < Level(α). We have

$$0 < (\alpha,\alpha) = \sum c_i(\alpha,\alpha_i).$$

Since all the c_i are positive, it assures the existence of a i_0 such that $(\alpha,\alpha_{i_0)} > 0$. Then from the lemma, $\beta = s_{\alpha_{i_0}}(\alpha) > 0$ ($\alpha \neq \alpha_{i_0}$ because $\text{Level}(\alpha) > 1$). The root β can be expanded as

$$\beta = \sum_{j\neq i_0} c_j\alpha_j + \left(c_{i_0} - \sum_{j\neq i_0}\frac{c_j}{|\alpha_{i_0}|^2}(\alpha,\alpha_{i_0})\right)\alpha_{i_0}. \tag{3.294}$$

Since $(\alpha,\alpha_{i_0}) > 0$, it implies $\text{Level}(\beta) < \text{Level}(\alpha)$ and thus $\beta = s'\alpha_j$ for a certain $s' \in W'$. So $\alpha = s_{\alpha_{i_0}}s'\alpha_j$ with $s_{\alpha_{i_0}}s' \in W'$. This conclude the induction. For $\alpha < 0$, the same result holds by writing $-\alpha = s\alpha_j$ and $\alpha = ss_{\alpha_j}\alpha_j$.

Now it remains to prove that $W' \subseteq W$. For a $\alpha \in \Phi$, we write $\alpha = s\alpha_j$ with $s \in W'$. Then

$$s_\alpha = ss_{\alpha_j}s^{-1} \in W'.$$

□

3.9.2 Dynkin diagram

Proposition 3.142.
If α and β are simple roots, then the angle $\theta_{\alpha,\beta}$ can only take the values 90°, 120°, 135° or 150°.

Proof. No proof. □

In order to draw the **Dynkin diagram** of a Lie algebra, one draws a circle for each simple root, and one joins the roots with 1, 2 or 3 lines, following that the value of the angle is 120°, 135° or 150°. If the roots are orthogonal (angle 90°), they are not connected. If the length of a root is maximal, the circle is left empty. If not, it is filled.

One easily determines the number of lines between two roots by the following proposition.

Proposition 3.143.
If α and β are two simple roots with $(\alpha,\alpha) \leqslant (\beta,\beta)$, then

$$\frac{(\alpha,\alpha)}{(\beta,\beta)} = \begin{cases} 1 & \text{if } \theta_{\alpha,\beta} = 120° \\ 2 & \text{if } \theta_{\alpha,\beta} = 135° \\ 3 & \text{if } \theta_{\alpha,\beta} = 150°. \end{cases} \tag{3.295}$$

Proof. No proof. □

If M is a weight of a representation, its **Dynkin coefficients** are

$$M_i = \frac{2(M,\alpha_i)}{(\alpha_i,\alpha_i)}, \tag{3.296}$$

and we can compute the Dynkin coefficients from one weight to another by the simple formula

$$(M - \alpha_j)_i = M_i - A_{ij}. \tag{3.297}$$

A weight is **dominant** if all its Dynkin coefficients are strictly positive.

3.9.2.1 Strings of roots

Let α, β be two roots with respect to $\mathfrak{h}$ and suppose $\beta \neq 0$. We denote by α^β the largest integer m such that $\alpha + m\beta$ is a root and by α_β the one such that $\alpha - m\beta$ is a root. Let $x \in \mathfrak{g}_\alpha$; since the Killing form is nondegenerate, there exists a $y \in \mathfrak{g}$ such that $B(x, y) \neq 0$. Using the root space decomposition (3.443) for y and corollary 3.79, $B(x, y) = B(x, y_{-\alpha})$. Then

$$\forall x \in \mathfrak{g}_\alpha, \exists y \in \mathfrak{g}_{-\alpha} \text{ such that } B(x, y) \neq 0.$$

In particular if α is a root, $-\alpha$ is also a root and the restriction of B to $\mathfrak{h} \times \mathfrak{h}$ is nondegenerate because $\mathfrak{h} = \mathfrak{g}_0$. So

$$\forall \mu \in \mathfrak{h}^*, \exists! h_\mu \in \mathfrak{h} \text{ such that } \forall h \in \mathfrak{g}, B(h, h_\mu) = \mu(h).$$

This is a general result about nondegenerate (here we use the semi-simplicity assumption) bilinear forms on a vector space. If $B(x, y) = B_{ij}x^i y^j$ and $a(x) = a_i x^i$, then a vector v such that $B(x, v) = a(x)$ exists, is unique and is given by coordinates $v^k = B^{ki}a_i$ where the matrix (B^{ij}) is the inverse of (B_{ij}).

We will sometimes use the following notation if α and β are roots:

$$(\alpha, \beta) = B(h_\alpha, h_\beta), \qquad |\alpha|^2 = (\alpha, \alpha).$$

By proposition 3.151, the roots come by pairs $(\alpha, -\alpha)$. For each of them, we choose $x_\alpha \in \mathfrak{g}_\alpha$. Our choice of $x_{-\alpha}$ is made as following. From discussion at page 93 we can find a $x_{-\alpha} \in \mathfrak{g}_{-\alpha}$ such that $B(x_{-\alpha}, x_\alpha) = 1$. Note that this choice is unambigous: if we had chosen first $x_{-\alpha} \in \mathfrak{g}_{-\alpha}$, this construction would have given the same x_α than our starting point. Note also that $h_{-\alpha} = -h_\alpha$. These x_α fulfil $[x_\alpha, x_{-\alpha}] = h_\alpha$.

Problem and misunderstanding 14.
Here the notation Δ does not follow our convention of subsection 3.8.1.3.

Let Δ be the set of non zero roots. We define an antisymmetric map $c \colon \Delta \times \Delta \to \mathbb{C}$ as following. If $\alpha, \beta \in S$ are such that $\alpha + \beta \notin \Delta$, we pose $c(\alpha, \beta) = 0$. If $\alpha + \beta \in \Delta$,

$$[x_\alpha, x_\beta] = c(\alpha, \beta)x_{\alpha+\beta}. \tag{3.298}$$

It is easy to see that $c(\alpha, \beta) = -c(\beta, \alpha)$.

Proposition 3.144.
If $\alpha, \beta, \alpha + \beta \in \Delta$, then

(i)
$$c(-\alpha, \alpha + \beta) = c(\alpha + \beta, -\beta) = c(-\beta, -\alpha),$$

(ii) If $\alpha, \beta, \gamma, \delta \in \Delta$ and $\alpha + \beta + \gamma + \delta = 0$ wile δ is neither $-\alpha$, nor $-\beta$ nor $-\gamma$, then
$$c(\alpha, \beta)c(\gamma, \delta) + c(\beta, \gamma)c(\alpha, \delta) + c(\gamma, \alpha)c(\beta, \delta) = 0, \tag{3.299}$$

(iii) if $\beta \neq \alpha \neq -\beta$, then
$$c(\alpha, \beta) + c(-\alpha, -\beta) = c(\alpha, -\beta)c(-\alpha, \beta) - B(h_\alpha, h_\beta),$$

(iv) if $\alpha + \beta \neq 0$ then
$$2c(\alpha, \beta)c(-\alpha, -\beta) = \beta^\alpha(1 + \beta_\alpha)\alpha(h_\alpha). \tag{3.300}$$

Proof. From our choice of x_α, we find that $B(x_\beta, x_{-\beta}) = B(x_{-\alpha}, x_\alpha) = B(x_{\alpha+\beta}, x_{\alpha-\beta}) = 1$, but

$$\begin{aligned} B\big(c(-\alpha,\alpha+\beta)x_\beta, x_{-\beta}\big) &= B\big(x_{-\alpha}, c(\alpha+\beta,-\beta)x_\alpha\big) \\ &= B\big(x_{\alpha+\beta}, c(-\beta,-\alpha)x_{-\alpha-\beta}\big). \end{aligned} \tag{3.301}$$

This proves (i). In order to prove (ii), suppose that

$$c(\alpha,\beta)c(\gamma,\delta) = B\Big(\big[[x_\alpha,x_\beta],x_\gamma\big],x_\delta\Big) \tag{3.302}$$

Then the Jacobi identity gives the result:

$$\begin{aligned} 0 &= B\Big(\big[[x_\alpha,x_\beta],x_\gamma\big],x_\delta\Big) + B\Big(\big[[x_\beta,x_\gamma],x_\alpha\big],x_\delta\Big) + B\Big(\big[[x_\gamma,x_\alpha],x_\beta\big],x_\delta\Big) \\ &= c(\alpha,\beta)c(\gamma,\delta) + c(\beta,\gamma)c(\alpha,\delta) + c(\gamma,\alpha)c(\beta,\delta), \end{aligned} \tag{3.303}$$

Here, we used the hypothesis $-\gamma \neq \delta \neq -\beta$ by supposing that (3.302) still hold after permutation of α,β,γ. Now we show the (3.302) is true. The assumptions imply $\alpha+\beta = -(\gamma+\delta) \neq 0$, then

$$\begin{aligned} B\big(\big[[x_\alpha,x_\beta],x_\gamma\big],x_\delta\big) &= B\big([x_\alpha,x_\beta],[x_\gamma,x_\delta]\big) \\ &= c(\alpha,\beta)c(\gamma,\delta)B(x_{\alpha+\beta},x_{\gamma+\delta}) \\ &= c(\alpha,\beta)c(\gamma,\delta). \end{aligned} \tag{3.304}$$

Now we turn our attention to (iii). If α and β fulfil the condition $\beta \neq \alpha \neq -\beta$, we can apply (ii) on the quadruple $(\alpha,\beta,-\alpha,-\beta)$ to get $c(\alpha,\beta)c(-\alpha,-\beta) = -B\big([x_\alpha,x_\beta],[x_{-\alpha},x_{-\beta}]\big)$. If we replace β by $-\beta$ and if we make the difference between the two expressions,

$$\begin{aligned} c(\alpha,\beta)c(-\alpha,-\beta) &= -B\big([x_\alpha,x_\beta],[x_{-\alpha},x_{-\beta}]\big) + B\big([x_\alpha,x_{-\beta}],[x_{-\alpha},x_\beta]\big) \\ &= B\big([x_\alpha,[x_{-\beta},x_{-\alpha}]],x_\beta\big) - B\big([x_{-\alpha},[x_\alpha,x_{-\beta}]],x_\beta\big) \\ &= -B\big([x_{-\alpha},x_\alpha],[x_{-\beta},x_\beta]\big) \\ &= -B(h_\alpha,h_\beta). \end{aligned} \tag{3.305}$$

In order to prove (iv), we consider $\alpha+\beta \neq 0$ and we pose

$$d(\alpha,\beta) = c(\alpha,\beta)c(-\alpha,-\beta) - \frac{1}{2}\beta^\alpha(1+\beta_\alpha)\alpha(h_\alpha).$$

Our aim is to prove that it is zero. We will do it by induction on β^α. First $\beta^\alpha = 0$ means that $\beta+\alpha = 0$, so that $c(\alpha,\beta) = 0$ and $d(\alpha,\beta) = 0$. Now we suppose that $\beta^\alpha > 0$ and that (iv) is yet checked for lower cases. Note that $\beta+\alpha \in \Delta$ and $(\beta+\alpha)+\alpha \neq 0$ because -2α is not a root. Then $\beta = 2\alpha$ is not possible. From the fact that $(\beta+\alpha)^\alpha = \beta^\alpha - 1$, we conclude $d(\alpha,\beta+\alpha) = 0$. Then

$$c(\alpha,\alpha+\beta)c(-\alpha,-\alpha-\beta) = c(\alpha,-\alpha-\beta)c(-\alpha,\alpha+\beta) - B(h_\alpha,h_{\alpha+\beta}).$$

On the other hand, (i) and the antisymmetry of c give

$$c(-\alpha,\alpha+\beta) = c(-\beta,-\alpha) = -c(-\alpha,-\beta) \tag{3.306a}$$

and

$$c(\alpha,-\alpha-\beta) = c(\beta,\alpha) = -c(\alpha,\beta) \tag{3.306b}$$

With all this

$$\begin{aligned} d(\alpha,\beta+\alpha) &= c(\alpha,\alpha+\beta)c(-\alpha,-\alpha-\beta) - \frac{1}{2}(\alpha+\beta)^\alpha(1+(\alpha+\beta)_\alpha)\alpha(h_\alpha) \\ &= c(\alpha,\beta)c(-\alpha,-\beta) - k(\alpha,\beta) \end{aligned} \tag{3.307}$$

where $k(\alpha,\beta) = B(h_\alpha,h_{\alpha+\beta}) + \frac{1}{2}(\alpha+\beta)^\alpha(1+(\alpha+\beta)_\alpha)\alpha(h_\alpha)$. But $h_{\alpha+\beta}$ is definied in order to have $B(h,h_{\alpha+\beta}) = (\alpha+\beta)(h)$ for any $h \in \mathfrak{h}$. Then using $2\beta(h_\alpha) = (\beta_\alpha - \beta^\alpha)\alpha(h_\alpha)$, we find $k(\alpha,\beta) = \frac{1}{2}\alpha(h_\alpha)\beta^\alpha(1+\beta_\alpha)$. □

Proposition 3.145.
Let

$$\mathfrak{h}_\mathbb{R} = \sum_{\alpha\in\Delta} \mathbb{R}h_\alpha. \tag{3.308}$$

(i) *Any root is real on* $\mathfrak{h}_{\mathbb{R}}$*,*

(ii) *the Killing form is real and strictly positive definite on* $\mathfrak{h}_{\mathbb{R}}$*,*

(iii) $\mathfrak{h} = \mathfrak{h}_{\mathbb{R}} \oplus i\mathfrak{h}_{\mathbb{R}}$.

The last item shows that $\mathfrak{h}_{\mathbb{R}}$ is a real form of $\mathfrak{h}$. Remark also that $\mathfrak{h}_{\mathbb{R}}$ can also be written as

$$\mathfrak{h}_{\mathbb{R}} = \{h \in \mathfrak{h} \text{ st } \alpha(h) \in \mathbb{R}\, \forall\alpha \in \Phi\}.$$

Proof. Let $\beta \in \Delta$; we looks at $\beta(h_\alpha)$. From (ii) of theorem 3.151, we know that $\alpha(h_\alpha)$ is real and positive, and (iii) makes $\beta(h_\alpha)$ real. From the formula $B(h_\alpha, h_\beta) = \sum_{\gamma\in\Delta} \gamma(h_\alpha)\gamma(h_\beta)$, the Killing form is real and positive definite on $\mathfrak{h}_{\mathbb{R}} \times \mathfrak{h}_{\mathbb{R}}$. If $B(h, h) = 0$ for a certain $h \in \mathfrak{h}_{\mathbb{R}}$, we find $\alpha(h) = 0$ for all $\alpha \in \Delta$. Then any $x = x^\alpha x_\alpha \in \mathfrak{g}$ commutes with h because

$$[h, x] = \sum_{\alpha\in\Phi} a^\alpha(\operatorname{ad} h)x_\alpha = \sum_\alpha a^\alpha \alpha(h) = 0.$$

So h is in the center of $\mathfrak{g}$ and so $h = 0$ be cause $\mathfrak{g}$ is semisimple. Thus the Killing form is strictly positive definite on $\mathfrak{h}_{\mathbb{R}} \times \mathfrak{h}_{\mathbb{R}}$.

Now we are going to show that $\mathfrak{h} = \mathfrak{h}_{\mathbb{R}} \oplus i\mathfrak{h}_{\mathbb{R}}$. If $h \in \mathfrak{h}_{\mathbb{R}} \cap i\mathfrak{h}_{\mathbb{R}}$, it can be written as $h = ih'$ with $h, h' \in \mathfrak{h}_{\mathbb{R}}$. Then

$$0 < B(h, h) = B(ih', ih') = -B(h', h') < 0,$$

so that $h = 0$ because B is nondegenerate. This shows that $\mathfrak{h}_{\mathbb{R}} \cap i\mathfrak{h}_{\mathbb{R}} = 0$. It is clear that $\sum_{\alpha\in\Delta} \mathbb{C}h_\alpha \subset \mathfrak{h}$; thus it remains to be proved that $\mathfrak{h} \subset \sum_{\alpha\in\Delta} \mathbb{C}h_\alpha$. It it is not, we can build a linear function $\lambda\colon \mathfrak{h} \to \mathbb{C}$ which is not identically zero but which is zero on the subspace $\sum_{\alpha\in\Delta} \mathbb{C}h_\alpha$. Then there exists (only one) $h_\lambda \in \mathfrak{h}$ such that $B(h, h_\lambda) = \lambda(h)$ for every $h \in \mathfrak{h}$. In particular, $\alpha(h_\lambda) = 0$ for every $\alpha \in \Delta$ because $\alpha(h_\lambda) = B(h_\alpha, h_\lambda) = \lambda(h_\alpha)$. This implies that $h_\lambda = 0$, so that $\lambda \equiv 0$.

□

One interest in the third point of this proposition is that we are now able to see Δ as a subset of $\mathfrak{h}_{\mathbb{R}}^*$ because the definition of $\alpha \in \Delta$ on $\mathfrak{h}_{\mathbb{R}}$ only is sufficient to define α on the whole $\mathfrak{h}$.

If $\{e_i\}$ is a basis of a vector space V, we say that $x = x^i e_i > y = y^i e_i$ if $x - y = a^i e_i$ and the first non zero a^i is positive. This is the **lexicographic order** on V. It is clear that it doesn't works on a complex vector space (because in this case we should first define $a^i > 0$), but we can anyway get an order on Δ by seeing it as a subset of $\mathfrak{h}_{\mathbb{R}}$.

The following important result is the fact that a complex semisimple Lie algebra is determined by its root system.

Theorem 3.146.
Let $\mathfrak{g}$ *and* $\mathfrak{g}'$ *be two semisimple complex Lie algebras;* $\mathfrak{h}$ *and* $\mathfrak{h}'$*, Cartan subalgebras. We suppose that we have a bijection* $\Phi \to \Phi'$*,* $\alpha \to \alpha'$ *which preserve the root system:*

- $\alpha' + \beta' = 0$ *if and only if* $\alpha + \beta = 0$*,*
- $\alpha' + \beta'$ *is not a root if and only if* $\alpha + \beta$ *is also not a root,*
- $(\alpha + \beta)' = \alpha' + \beta'$ *whenever* $\alpha + \beta$ *is a root.*

Then we have a Lie algebra isomorphism $\eta\colon \mathfrak{g} \to \mathfrak{g}'$ *such that* $\eta(\mathfrak{h}) = \mathfrak{h}'$ *and* $\alpha' \circ \eta|_{\mathfrak{h}} = \alpha$*.*

Proof. From the assumptions, $\beta^\alpha = (\beta')^{\alpha'}$ and $\beta_\alpha = (\beta')_{\alpha'}$ and the point (ii) of theorem 3.151 makes $\alpha'(h_{\alpha'}) = \alpha(h_\alpha)$. The fourth point of the same theorem then gives

$$\beta'(h\alpha') = \beta(h_\alpha). \tag{3.309}$$

Now we choose a maximally linearly independent set $(\alpha_1, \ldots, \alpha_R)$ of roots of $\mathfrak{g}$. Because of theorem 3.150, this is a basis of $\mathfrak{h}^*$. For notational convenience, we put $h_r = h_{\alpha_r}$ and naturally, $h'_r = h_{\alpha'_r}$. It is easy to see that the set of h_r is a basis of $\mathfrak{h}$. Indeed if $a^r h_r = 0$ (with sum over r), then $B(h, a^r, h_r) = a^r\alpha_r(h) = 0$ which implies that $a^r\alpha_r|_{\mathfrak{h}} = 0$ but it is impossible because the α_r are free in $\mathfrak{h}^*$.

$$\begin{aligned}\{\alpha_1, \ldots, \alpha_r\} &\text{ is a basis of } \mathfrak{h}^*,\\ \{h_1, \ldots, h_r\} &\text{ is a basis of } \mathfrak{h}.\end{aligned}$$

Then the matrix $(A_{ij}) = \alpha_i(h_j)$ has non zero determinant. Since $\alpha'_i(h'_j) = \alpha_i(h_j)$, the set $\{\alpha'_1, \ldots, \alpha'_r\}$ is free and $\{h'_1, \ldots, h'_r\}$ is a basis of $\mathfrak{h}'$.

$$\begin{aligned}\{\alpha'_1, \ldots, \alpha'_r\} &\text{ is a basis of } \mathfrak{h}'^*,\\ \{h'_1, \ldots, h'_r\} &\text{ is a basis of } \mathfrak{h}'.\end{aligned}$$

Then can define an isomorphism $\eta_{\mathfrak{h}}\colon \mathfrak{h} \to \mathfrak{h}'$ by $\eta_{\mathfrak{h}}(h_i) = h'_i$. If $x \in \mathfrak{h}$ is decomposed as $x = a^r h_r$, from equation (3.309) we have $(\alpha'_i \circ \eta_{\mathfrak{h}})(a^r h_r) = a^r \alpha'(h'_r) = \alpha_i(h_r)$. Then

$$\alpha'_i \circ \eta_{\mathfrak{h}} = \alpha_i.$$

Let $\alpha \in \Phi$; we can write $\alpha = c_i\alpha_i$ and $\alpha' = c'_i\alpha'_i$ (with a sum over i). We have

$$c_i\alpha_i(h_k) = \alpha(h_j) = \alpha'(h_j) = c'_i\alpha_i(h_j). \tag{3.310}$$

As the determinant of $(\alpha_i(h_j))$ is non zero, this implies $c_i = c'_i$, so that

$$\alpha' \circ \eta_{\mathfrak{h}} = \alpha \tag{3.311}$$

because $\alpha' \circ \eta_{\mathfrak{h}} = c'_i(\alpha'_i \circ \eta_{\mathfrak{h}}) = c_i\alpha_i = \alpha$. Now we "just" have to extend $\eta_{\mathfrak{h}}$ into a Lie algebra isomorphism $\eta\colon \mathfrak{g} \to \mathfrak{g}'$. As before for each $\alpha \in \Delta$ we choose $x_\alpha \in \mathfrak{g}_\alpha$ such that $B(x_\alpha, x_{-\alpha}) = -1$ and $[x_{-\alpha}, x_\alpha] = h_\alpha$. We naturally do the same for $x_{\alpha'} \in \mathfrak{g}'_{\alpha'}$. We also consider the function c as before: $[x_\alpha, x_\beta] = c(\alpha,\beta)x_{\alpha+\beta}$. Since $\mathfrak{h} = \mathfrak{g}_0$, these x_α form a basis of $\mathfrak{g}\ominus\mathfrak{h}$ and η can be defined by the date of $\eta(x_\alpha)$. We set $\eta(x_\alpha) = a_\alpha x_{\alpha'}$ (without sum).

The condition $\eta\big([x_\alpha, x_\beta] = [\eta(x_\alpha), \eta(x_\beta)]\big)$ gives

$$c(\alpha,\beta)a_{\alpha+\beta} = c(\alpha',\beta')a_\alpha a_\beta \qquad \text{if } \alpha+\beta \neq 0 \tag{3.312a}$$

and

$$a_\alpha a_{-\alpha} = 1 \qquad \forall \alpha \in \Phi. \tag{3.312b}$$

These two conditions are necessary and also sufficient. Indeed there are three cases of $[x,y]$ to check: x, $y \in \mathfrak{h}$, one of these two is out of $\mathfrak{h}$ or x, y are booth out of $\mathfrak{h}$. In the third case, using (3.312a),

$$\eta([x_\alpha, x_\beta]) = c(\alpha,\beta)a_{\alpha+\beta}x_{\alpha'+\beta'} = x(\alpha',\beta')a_\alpha a_\beta x_{\alpha'+\beta'} = a_\alpha a_\beta[x_{\alpha'}, a_{\beta'}] = [\eta(x_\alpha), \eta(x_\beta)]. \tag{3.313}$$

If x, $y \in \mathfrak{h}$, then from theorem 3.150, $\eta([x,y]) = 0 = [\eta(x), \eta(y)]$. Using the fact that $[h, x_\alpha] = \alpha(h)x_\alpha$, we find the third case:

$$\eta\big([h_\beta, x_\alpha]\big) = \eta\big(\alpha(h_\alpha)x_\alpha\big) = \eta\big(\alpha'(h_{\beta'})x_\alpha\big) = a_\alpha[h_{\beta'}, x_{\alpha'}] = [\eta(h_\beta), \eta(x_\alpha)]. \tag{3.314}$$

Now we are going to find some $a_\alpha \in \mathbb{C}$ such that

- $a_\alpha a_{-\alpha} = 1$ for any α,
- $c(\alpha,\beta)a_{\alpha+\beta} = c(\alpha',\beta')a_\alpha a_\beta$ if $\alpha+\beta \neq 0$.

We consider the lexicagraphic order on Φ: this is the order on Φ seen as a subset of $\mathfrak{h}_{\mathbb{R}}$ on which we put the lexicagraphic order. For a root $\alpha > 0$, we will fix the coefficient a_α by an induction with respect to the order and put $a_{-\alpha} = a_\alpha^{-1}$. Let us consider $\rho > 0$ and suppose that a_α is already defined for $-\rho < \alpha < \rho$ in such a manner that $a_\alpha a_{-\alpha} = 1$ and $c(\alpha,\beta)a_{\alpha+\beta} = c(\alpha',\beta')a_\alpha a_\beta$ for every α, β such that α, β and $\alpha+\beta$ are stricly between $-\rho$ and ρ. We have to define a_ρ in such a way that if $a_{-\rho} = a_\rho^{-1}$, the second condition holds for every α, β such that α, β and $\alpha+\beta$ are no zero roots between $-\rho$ and ρ.

If such a pair (α,β) doesn't exist, there are no problem to put $a_\rho = a_{-\rho} = 1$. Let us suppose that such a pair exists: $\alpha+\beta = \rho$. Then $\beta^\alpha \neq 0$ and the point (iii) of proposition 3.144 shows that $c(\alpha,\beta) \neq 0$; in the same way, $(\beta')^{\alpha'} = \beta^\alpha \neq 0$ implies $c(\alpha',\beta') \neq 0$. We define

$$a_\rho = c(\alpha,\beta)^{-1}c(\alpha',\beta')a_\alpha a_\beta, \tag{3.315a}$$

$$a_{-\rho} = a_\rho^{-1}. \tag{3.315b}$$

Since the value of the right hand side of (3.300) doesn't change under $\alpha \to \alpha'$ and $\beta \to \beta'$, it gives $c(\alpha,\beta)c(-\alpha,-\beta) = c(\alpha',\beta')c(-\alpha',-\beta')$ and thus

$$\begin{aligned}c(-\alpha,-\beta)a_{-\rho} &= c(-\alpha,-\beta)c(\alpha,\beta)c(\alpha',\beta')^{-1}a_{-\alpha}a_{-\beta}\\ &= c(\alpha',\beta')c(-\alpha',-\beta')c(\alpha',\beta')^{-1}a_\alpha a_{-\beta}\\ &= c(-\alpha',-\beta')a_{-\alpha}a_{-\beta}.\end{aligned} \tag{3.316}$$

Thus the definition (3.315) fulfils the requirements for the pair (α, β). It should be shown whether that works as well with another pair (γ, δ) such that $-\rho \leqslant \gamma, \delta \leqslant \rho$ and $\gamma + \delta = \rho$. If this second pair is really different that (α, β), then δ is neither α nor β; it is also clear that δ is not $-\gamma$. Then formula (3.299) works with the quadruple $(-\alpha, -\beta, \gamma, \delta)$:

$$c(-\alpha,-\beta)c(\gamma,\delta)+c(-\beta,\gamma)c(-\alpha,\delta)+c(\gamma,-\alpha)c(-\beta,\delta)=0. \tag{3.317}$$

If $\alpha < 0$, the assumption $\alpha+\beta=\rho$ makes $\beta > \rho$, which is in contradiction with $-\rho \leqslant \beta \leqslant \rho$. Then $\alpha, \beta, \gamma, \delta > 0$ and moreover, the difference of any two of them is strictly between $-\rho$ and ρ. Since $\delta-\alpha = -(\gamma-\beta)$, if $\gamma-\beta$ is a root, $\delta-\alpha$ is also a root and the induction hypothesis gives

$$c(\gamma,-\beta)a_{\gamma-\beta} = c(\gamma',-\beta')a_\gamma a_{-\beta}, \tag{3.318a}$$
$$c(-\alpha,\delta)a_{-\alpha+\delta} = c(-\alpha',\delta')a_{-\alpha}a_\delta. \tag{3.318b}$$

If we take for the convention $a_\mu = 1$ whenever μ is not a root, these relations still hold if $\gamma-\beta$ is not a root. In the same way,

$$c(\gamma,-\alpha)a_{\gamma-\alpha} = c(\gamma',-\alpha')a_{-\alpha}a_\gamma, \tag{3.319a}$$
$$c(-\beta,\delta)a_{-\beta+\delta} = c(-\beta',\delta')a_{-\beta}a_\delta. \tag{3.319b}$$

As $\delta-\alpha=-(\gamma-\beta)$, we have $a_{\delta-\alpha}a_{\gamma-\beta}=1$ and in the same way, $a_{\gamma-\alpha}a_{\delta-\beta}=1$. Taking it into account and multiplicating (3.318a) by (3.318b) and (3.319a) by (3.319a), we find:

$$c(-\beta,\gamma)c(-\alpha,\delta) = c(-\beta',\gamma')c(-\alpha',\delta')a_{-\alpha}a_{-\beta}a_\gamma a_\delta \tag{3.320a}$$
$$c(\gamma,-\alpha)c(-\beta,\delta) = c(\gamma',-\alpha')c(-\beta',\delta')a_{-\alpha}a_{-\beta}a_\gamma a_\delta. \tag{3.320b}$$

We can use it to rewrite equation (3.317). After multiplication by $a_\alpha a_\beta a_{-\gamma}a_{-\delta}$,

$$c(-\alpha,-\beta)c(\gamma,\delta)a_\alpha a_\beta a_{-\gamma}a_{-\delta}+c(-\beta',\gamma')c(-\alpha',\delta')+c(\gamma',-\alpha')c(-\beta',\delta')=0. \tag{3.321}$$

But equation (3.317) is also true for $(\alpha',\beta',\gamma',\delta')$ instead of $(\alpha,\beta,\gamma,\delta)$, so that the last two terms can be replaced by only one term to give

$$c(-\alpha,-\beta)c(\gamma,\delta)a_\alpha a_\beta a_{-\gamma}a_{-\delta}-c(-\alpha',-\beta')c(\gamma',\delta')=0.$$

Since the pair (α,β) fulfils $c(-\alpha,-\beta)a_{-\alpha-\beta}=c(-\alpha',-\beta')a_{-\alpha}a_{-\beta}$, using $\alpha+\beta=\gamma+\delta$, we find

$$c(\gamma,\delta)a_{\gamma+\delta}=c(\gamma',\delta')a_\gamma a_\delta.$$

□

Corollary 3.147.
The elements $x_\alpha \in \mathfrak{g}_\alpha$ can be chosen in order to satisfy

- $B(x_\alpha, x_{-\alpha}) = 1$,
- $[x_\alpha, x_{-\alpha}] = h_\alpha$,
- $c(\alpha,\beta) = c(-\alpha,-\beta)$.

These vectors $x_\alpha \in \mathfrak{g}_\alpha$ are called **root vectors**.

Proof. We consider the isomorphism $\alpha \to \alpha$ from Φ to Φ; by the theorem this induces an isomorphism $\eta\colon \mathfrak{g} \to \mathfrak{g}$ given by some constants c_α:

$$\eta(x_\alpha) = c_{-\alpha}x_{-\alpha}$$

without sum on α, because of course $\eta(x_\alpha) \in \mathfrak{g}_{-\alpha}$. We choose $ax_\alpha \in \mathbb{C}$ in such a way that

$$a_\alpha^2 = -c_{-\alpha} \tag{3.322a}$$
$$a_\alpha a_{-\alpha} = 1, \tag{3.322b}$$

and then we put $y_\alpha = a_\alpha x_\alpha$. It is immediate that $B(y_\alpha, y_{-\alpha}) = 1$, thus the redefinition $x_\alpha \to y_\alpha$ doesn't change the obtained relations. Acting on y_α, the isomorphism η gives

$$\eta(y_\alpha) = a_\alpha c_{-\alpha}x_{-\alpha} = -a_{-\alpha}x_{-\alpha} = -y_{-\alpha}. \tag{3.323}$$

If $\alpha, \beta, \alpha+\beta \in \Delta$, we naturally define $c'(\alpha,\beta)$ by

$$[y_\alpha, y_\beta] = c'(\alpha,\beta) y_{\alpha,\beta}.$$

Using the fact that η is a Lie algebra automorphism of $\mathfrak{g}$ we have:

$$-c'(\alpha,\beta)y_{-(\alpha+\beta)} = \eta\big(c'(\alpha,\beta)y_{\alpha+\beta}\big) = [-y_{-\alpha}, -y_{-\beta}] = c'(-\alpha,-\beta)y_{-(\alpha+\beta)}. \tag{3.324}$$

□

From now we always our x_α in this way.

Remark 3.148.
It is also possible to choice the x_α in such a way that

- $B(x_\alpha, x_{-\alpha}) = -1$*,*
- $c(\alpha,\beta) = c(-\alpha,-\beta)$*.*

This is the choice of the reference [17].

Here is a characterization for Cartan subalgebras of semisimple Lie algebras. This is sometimes taken as the *definition* of a Cartan subalgebra in books devoted to semisimple Lie algebras (for example in [3]).

Proposition 3.149.
A subalgebra $\mathfrak{h}$ of a semisimple Lie algebra $\mathfrak{g}$ is a Cartan subalgebra if and only if

- $\mathfrak{h}$ *is maximally abelian in* $\mathfrak{g}$*,*
- *the endomorphism* $\operatorname{ad} h$ *is semisimple for every* $h \in \mathfrak{h}$*.*

Here, "semisimple" means "diagonalisable", cf. definition at page 3.6.1.

Proof. Necessary condition. We know from theorem 3.150 that $\mathfrak{h}$ is abelian and from proposition 3.74 that it is maximally nilpotent. Then it is maximally abelian. On the other hand, let $h \in \mathfrak{h}$; the endomorphism $\operatorname{ad} h$ is diagonalisable with respect to the decomposition $\mathfrak{g} = \mathfrak{h} \bigoplus_{\alpha\in\Delta} \mathfrak{h}_\alpha$.
Sufficient condition. Firstly it is clear that a maximal abelian subalgebra is nilpotent and the $\operatorname{ad} h_i$ are simultaneously diagonalisable for the different $h_i \in \mathfrak{h}$. Let $\{x_1, \ldots, x_n\}$ be a basis of $\mathfrak{g}$ which diagonalise all the $\operatorname{ad} h_i$. In this basis, if $(\operatorname{ad} h)_{ii} = 0$ for any $h \in \mathfrak{h}$, then $x_i \in \mathfrak{h}$: if it was not, $\mathfrak{h} \cup \{x_i\}$ would be abelian.

Let $x \in \mathfrak{g}$ such that $(\operatorname{ad} h)x \in \mathfrak{h}$ for every $h \in \mathfrak{h}$. Suppose that x has a x_i-component with $x_i \notin \mathfrak{h}$. There is a $h \in \mathfrak{h}$ with $(\operatorname{ad} h)_{ii} \neq 0$. Then $(\operatorname{ad} h)x$ has a x_i-component and can't lies in $\mathfrak{h}$.

□

This characterization of Cartan subalgebras is used to prove the existence of Cartan subalgebra for any complex semisimple Lie algebra.

Theorem 3.150.
The Cartan algebra of a complex semisimple Lie algebra is abelian and the dual is spanned by the roots: $\operatorname{Span}\Phi = \mathfrak{h}^*$*.*

Proof. Let α be a non zero root; from the point (ii) of proposition 3.220, there exists a $v \in \mathfrak{g}_\alpha$ such that for any $x \in \mathfrak{h}$, $[x,v] = \alpha(x)v$. Since $\dim \mathfrak{g}_\alpha = 1$ it is in fact true for any $v \in \mathfrak{g}_\alpha$. In particular $\forall v \in \mathfrak{g}_\alpha$ and $h \in \mathfrak{h}$, we have $[h,x] = \alpha(h)x$.

Let $\mathfrak{n} \subset \mathfrak{h}$ be the set of elements which are annihilated by all the roots:

$$\mathfrak{n} = \{H \in \mathfrak{h} \text{ st } \alpha(H) = 0\,\forall\alpha \in \Phi\}. \tag{3.325}$$

First remark that

$$[\mathfrak{g}_\alpha, \mathfrak{n}] = 0 \tag{3.326}$$

because for $x \in \mathfrak{g}_\alpha$ and $h \in \mathfrak{n} \subset \mathfrak{h}$, we have $[h,x] = \alpha(h)x = 0$. An other property of $\mathfrak{n}$ is

$$[\mathfrak{h},\mathfrak{h}] \subset \mathfrak{n}. \tag{3.327}$$

Indeed consider a root α and $x \in \mathfrak{g}_\alpha$. We have

$$\begin{aligned} -\alpha([h,h'])x &= [x,[h,h']] = [h,[h',x]] + [h',[x,h]] = \alpha(h)[h',x] + \alpha(h')[x,h] \\ &= \alpha(h)\alpha(h') - \alpha(h')\alpha(h) = 0. \end{aligned} \tag{3.328}$$

If $x \in \mathfrak{g}$ is decomposed as $x = \sum_{\alpha\in\Phi} x_\alpha$ and if $n \in \mathfrak{n}$, then

$$[x, n] = \sum_\alpha [x_\alpha, n] = \sum_\alpha \alpha(n) x_\alpha = 0.$$

In particular, $\mathfrak{n}$ is an ideal[22]. Moreover, the fact that $\mathfrak{n} \subset \mathfrak{h}$ makes $\mathfrak{n}$ a *nilpotent* ideal in the semisimple Lie algebra $\mathfrak{g}$. Then $\mathfrak{n} = 0$. Equation (3.326) makes $\mathfrak{h}$ abelian while equation (3.327) says that no element of $\mathfrak{h}$ is anihilated by all the roots. This implies that $\operatorname{Span}\Phi = \mathfrak{h}^*$. To see it more precisely, if Φ don't span a certain (dual) basis element e_i^* of $\mathfrak{h}^*$, then a basis of $\operatorname{Span}\Phi$ is at most $\{e_j\}_{j\neq i}$. Then it is clear that $\alpha(e_i) = 0$ for any root α. □

Theorem 3.151.
If α, β are roots of a semisimple Lie algebra $\mathfrak{g}$ with respect to a Cartan subalgebra $\mathfrak{h}$, then

(i) if $x_\alpha \neq 0 \in \mathfrak{g}_\alpha$ fulfils $[h, x_\alpha] = \alpha(h)x_\alpha$ for all $h \in \mathfrak{h}$, then $\forall y \in \mathfrak{g}_{-\alpha}$

$$[x_\alpha, y] = B(x_\alpha, y)h_\alpha,$$

(ii) $\alpha(h_\alpha)$ is rational and positive. Moreover

$$\alpha(h_\alpha) \sum_{\gamma\in\Phi} (\gamma_\alpha - \gamma^\alpha)^2 = 4,$$

(iii) $2\beta(h_\alpha) = (\beta_\alpha - \beta^\alpha)\alpha(h_\alpha)$,

(iv) the forms $0, \alpha, -\alpha$ are the only integer multiples of α which are roots,

(v) $\dim \mathfrak{g}_\alpha$=1,

(vi) any k which makes $\beta + k\alpha$ a root lie between $-\beta_\alpha$ and β^α. In other words, $\beta + k\alpha \in \Phi$ is only true with $-\beta_\alpha \leqslant k \leqslant \beta^\alpha$.

Proof. The fact that $y \in \mathfrak{g}_{-\alpha}$ and that $x \in \mathfrak{g}_\alpha$ make $[x, y] \in \mathfrak{g}_0 = \mathfrak{h}$. Now we consider $h \in \mathfrak{h}$ and the invariance formula (3.21). We find:

$$B(h, [x_\alpha, y]) = -B([x_\alpha, h], y) = \alpha(h)B(x_\alpha, y) = B(h, h_\alpha)B(x_\alpha, y) = B(h, B(x_\alpha, y)h_\alpha). \tag{3.329}$$

Since it is true for any $h \in \mathfrak{h}$ and B is nondegenerate on $\mathfrak{h}$ we find the first point. In order to prove (ii), we consider

$$U = \bigoplus_{-\beta_\alpha \leqslant m \leqslant \beta^\alpha} \mathfrak{g}_{\beta+m\alpha}.$$

By definition of α_β and α^β, each term of the sum is a root space. If $z \in \mathfrak{g}_\alpha \oplus \mathfrak{g}_{-\alpha}$, then U is stable under $\operatorname{ad} z$ because the terms in $\operatorname{ad} zU$ are of the form $[z, x_{\beta+m\alpha}] \in \mathfrak{g}_{\beta+m\alpha\pm\alpha}$. Note however that this $\operatorname{ad} zU$ is not *equal* to U.

Let $x_\alpha \neq 0 \in \mathfrak{g}_\alpha$. There exists a $y \in \mathfrak{g}_{-\alpha}$ such that $[x_\alpha, y] = B(x_\alpha, y)h_\alpha$ (here we use semi-simplicity). By fitting the norm of y, we can choose it in order to get $[x_\alpha, y] = h_\alpha$, so that

$$\operatorname{ad} h_\alpha = [\operatorname{ad} x_\alpha, \operatorname{ad} y].$$

Now we look at the restriction of $\operatorname{ad} h_\alpha$ to U:

$$\operatorname{Tr}(\operatorname{ad} h_\alpha) = \operatorname{Tr}(\operatorname{ad} x_\alpha \circ \operatorname{ad} y) - \operatorname{Tr}(\operatorname{ad} y \circ \operatorname{ad} x_\alpha) = 0. \tag{3.330}$$

Since $h_\alpha \in \mathfrak{h} = \mathfrak{g}_0$, we have $\operatorname{ad} h_\alpha \colon U \to U$, so that the annihilation of the trace of $\operatorname{ad} h_\alpha$ can be particularised to

$$\operatorname{Tr}(\operatorname{ad} h_\alpha|_U) = 0.$$

On the other hand, by definition $\operatorname{ad} h_\alpha - (\beta + m\alpha)(h_\alpha)$ is nilpotent on $\mathfrak{g}_{\beta+m\alpha}$. Then it has a vanishing trace:

$$\sum_m \operatorname{Tr}(\operatorname{ad} h_\alpha - (\beta + m\alpha)h_\alpha) = 0.$$

[22] Ça me semble quand même fort de prouver que c'est le centralisateur pour dire que c'est un idéal. D'autant plus que je pourais directement dire que $\mathfrak{n}$ est centralisateur dans un semisimple et donc nulle.

But we had yet seen that the term with $\operatorname{ad} h_\alpha$ is zero; then

$$\sum_{-\beta_\alpha\leqslant m\leqslant\beta^\alpha}(\beta+m\alpha)h_\alpha\dim\mathfrak{g}_{\beta+m\alpha}=0. \tag{3.331}$$

If we suppose that $\alpha(h_\alpha)=0$ this gives $\beta(h_\alpha)=0$. Since this conclusion is true for any root β, we find $B(h,h_\alpha)=0$ for any $h\in\mathfrak{h}$. In other words, $\alpha(h)=0$ for any $h\in\mathfrak{h}$. This contradicts the assumption, so that we conclude $\alpha(h_\alpha)\neq 0$.

Let $V=\mathfrak{h}+(x_\alpha)+\sum_{m<0}\mathfrak{g}_{m\alpha}$ where (x_α) is the one dimensional space spanned by x_α. On the one hand, from the definition of x_α, $\operatorname{ad}x_\alpha\mathfrak{h}\subset(x_\alpha)$ and $\operatorname{ad}x_\alpha\mathfrak{g}_{m\alpha}\subset\mathfrak{g}_{(m+1)\alpha}$. On the other hand, $y\in\mathfrak{g}_{-\alpha}$ is defined by the relation $[x_\alpha,y]=h_\alpha$, then $\operatorname{ad}y\mathfrak{h}\subset\mathfrak{g}_{-\alpha}\subset\sum_{m<0}\mathfrak{g}_{m\alpha}$, $\operatorname{ad}y(x_\alpha)\subset\mathfrak{g}_0=\mathfrak{h}$ and $\operatorname{ad}y\sum_{m<0}\mathfrak{g}_{m\alpha}=\sum_{m<0}\mathfrak{g}_{(m-1)\alpha}$. All this make V invariant under $\operatorname{ad}x_\alpha$ and $\operatorname{ad}y$.

Since $\operatorname{ad}h_\alpha=[\operatorname{ad}x_\alpha,\operatorname{ad}y]$, the trace of $\operatorname{ad}h_\alpha$ is zero so that the invariance of V gives

$$\operatorname{Tr}(\operatorname{ad}h_\alpha|_V)=0.$$

By the definition of x_α particularised to $h\to h_\alpha$, we have $\operatorname{Tr}(\operatorname{ad}h_\alpha|_{(x_\alpha)})=\alpha(h_\alpha)$. By the definition of $\mathfrak{g}_0$, for any $x\in\mathfrak{h}$ and $v\in\mathfrak{g}_0$, $\operatorname{ad}x$ is nilpotent on v. Taking h_α as x, we see that $(\operatorname{ad}h_\alpha)h$ don't contain "h-component". Then $\operatorname{Tr}(\operatorname{ad}h_\alpha|_{\mathfrak{h}})=0$. Finally the operator $(\operatorname{ad}h_\alpha-m\alpha(h_\alpha))$ is nilpotent on $\mathfrak{g}_{m\alpha}$, so that $\operatorname{Tr}(\operatorname{ad}h_\alpha|_{\mathfrak{g}_{m\alpha}})=\operatorname{Tr}(m\alpha(\alpha)|_{\mathfrak{g}_{m\alpha}})=m\alpha(h_\alpha)\dim\mathfrak{g}_{m\alpha}$. All this gives

$$\alpha(h_\alpha)\left(1+\sum_{m<0}m\dim\mathfrak{g}_{m\alpha}\right)=0. \tag{3.332}$$

As we saw that $\alpha(h_\alpha)\neq 0$, we conclude that $\dim\mathfrak{g}_{m\alpha}=0$ for $m<-1$ and $\dim\mathfrak{g}_{-\alpha}=1$. This proves (v).

This also prove (iv) in the particular case of *integer* multiples. It is rather simple to get relations such that $0_\alpha=1$, $0^\alpha=1$, $\alpha_\alpha=2$, $(-\alpha)_\alpha=0$, and it is easy to check (iii) in the cases $\beta=-\alpha,0,\alpha$. Now we turn our attention to the case in which β is not an integer multiple of α. By (iv) applied to $\alpha\to\beta+m\alpha$, we have $\dim\mathfrak{g}_{\beta+m\alpha}=1$ whenever $-\beta_\alpha\leqslant m\leqslant\beta^\alpha$.

From equation (3.331), $\sum_{-\beta_\alpha\leqslant m\leqslant\beta^\alpha}(\beta(h_\alpha)+m\alpha(h_\alpha))=0$, then

$$(\beta_\alpha+\beta^\alpha+1)\beta(h_\alpha)=(\sum_m m)\alpha(h_\alpha)=\left(\frac{\beta_\alpha(\beta_\alpha+1)}{2}-\frac{(\beta^\alpha-1)\beta^\alpha}{2}\right)\alpha(h_\alpha). \tag{3.333}$$

This gives (iii). Now we consider the formula of theorem 3.152 in the case $x=y=h_\alpha$ and we use the fact that $B(h,h_\alpha)=\alpha(h)$ in the case $h=h_\alpha$:

$$B(h_\alpha,h_\alpha)=\alpha(h_\alpha)=\sum_{\gamma\in\Phi}\dim\mathfrak{g}_\gamma\gamma(h_\alpha)^2=\sum_{\gamma\in\Phi}\gamma(h_\alpha)^2. \tag{3.334}$$

Since $\beta(h_\alpha)=\frac{1}{2}(\beta_\alpha-\beta^\alpha)\alpha(h_\alpha)$, we find (ii). In order to prove (iv), we consider $\beta=c\alpha$ for a $c\in\mathbb{C}$. By (iii), $2c\alpha(h_\alpha)=(\beta_\alpha-\beta^\alpha)\alpha(h_\alpha)$, so that c is an half integer: $c=p/2$ with $p\in\mathbb{Z}$. If c is non zero, we can interchange α and β and see that $\alpha=c^{-1}\beta$ implies $c^{-1}=q/2$ with $q\in\mathbb{Z}$. It is clear the $pq=4$. But we had already discussed the case of integer multiples of α, so that we can suppose that p is odd. The only odd p such that $pq=3$ with $q\in\mathbb{Z}$ are $p=1,-1$, which are two excluded cases: they are $\alpha=\pm 2\beta$ which lies in the case of integer multiples.

It remains to prove (vi). By definition of β^α, the form $\beta+(\beta^\alpha+1)\alpha$ is not a root. But it remains possible that $\beta+(\beta^\alpha+2)\alpha$ is. We suppose that $k_1,\ldots,k_p$ are the p positive integers such that $\beta+k_i\alpha\in\Phi$. We pose

$$W=\bigoplus_{i=1}^p\mathfrak{g}_{\beta+k_i\alpha}.$$

As usual we see that W is stable under $\operatorname{ad}x_\alpha$ and $\operatorname{ad}y$ (because $k_i\geqslant\beta^\alpha+2$). The trace of $\operatorname{ad}g_\alpha$ on W is zero, thus

$$0=\sum_{i=1}^p(\beta+k_i\alpha)(h_\alpha). \tag{3.335}$$

By (iii), we find

$$p(\beta_\alpha-\beta^\alpha)\alpha(h_\alpha)=2(k_1+\ldots+k_p)>p(\beta^\alpha+1).$$

This is not possible because it would gives $-\beta^\alpha-\beta_\alpha>2$. $\square$

Theorem 3.152.
Let $\mathfrak{g}$ be a Lie algebra, $\mathfrak{h}$ a Cartan subalgebra of $\mathfrak{g}$ and B the Killing form of $\mathfrak{g}$. Then for all x, $y \in \mathfrak{h}$,

$$B(x,y) = \sum_{\gamma\in\Phi} d_\gamma \gamma(x)\gamma(y) \tag{3.336}$$

where $g_\gamma = \dim \mathfrak{g}_\gamma$.

Proof. We are seeing $\mathfrak{g}$ as a $\mathfrak{h}$-module for the adjoint representation. In particular, proposition 3.220 makes $\mathfrak{g}$ a direct sum of the $\mathfrak{h}$-submodules $\mathfrak{g}_\gamma$. Then

$$B(x,y) = \operatorname{Tr}(\operatorname{ad} x^2) = \sum_{\gamma\in\Phi} \operatorname{Tr}(\operatorname{ad} x|_\gamma^2) \tag{3.337}$$

where $\operatorname{ad} x|_\gamma$ means the restriction of $\operatorname{ad} x$ to $\mathfrak{g}_\gamma$. It is clear that $\operatorname{ad} x|_\gamma - \gamma(x)$ is nilpotent, then $\operatorname{ad} x|_\gamma^2 - \gamma(x)^2$ is also nilpotent because

$$\operatorname{ad} x|_\gamma^2 - \gamma(x)^2 = (\operatorname{ad} x|_\gamma + \gamma(x))(\operatorname{ad} x|_\gamma - \gamma(x))$$

and the fact that these two terms commute. The trace of a nilpotent endomorphism is zero, then $\operatorname{Tr}(\operatorname{ad} x|_\gamma^2 - \gamma(x)^2) = 0$ or for all $x \in \mathfrak{g}$,

$$B(x,x) = \sum_{\gamma\in\Phi} d_\gamma \gamma(x)^2. \tag{3.338}$$

on the other hand, we know that a quadratic form determines only one bilinear form. Here the form (3.338) gives

$$B(x,y) = \sum_{\gamma\in\Phi} d_\gamma \gamma(x)\gamma(y).$$

□

3.9.3 Weyl: other results

Proposition 3.153.
Two immediate properties of the Weyl group are

(i) W is a finite group of orthogonal transformations of V,

(ii) if r is an orthogonal transformation of V, the $s_{r\alpha} = rs_\alpha r^{-1}$.

Proof. First item. By definition of an abstract root system, W leaves Δ invariant; since V is spanned by V, it implies that W also leaves V invariant. From an easy computation, $(s_\alpha\varphi, s_\alpha\phi) = (\varphi,\phi)$. Since Δ is a finite set, there are only a finite number of common permutations of elements of Δ *a fortiori* W is finite.
Second item. It is easy to see that $s_{r\alpha}(r\varphi) = rs_\alpha\varphi$, then $s_{r\alpha} = r\circ s_\alpha \circ r^{-1}$. □

We introduce the **root reflexion** $s_\alpha \colon \mathfrak{h}_\mathbb{R}^* \to \mathfrak{h}_\mathbb{R}^*$ for $\alpha \in \Phi$ and $\varphi \in \mathfrak{h}_\mathbb{R}^*$ by

$$s_\alpha(\varphi) = \varphi - \frac{2(\varphi,\alpha)}{|\alpha|^2}\alpha. \tag{3.339}$$

Proposition 3.154.
If $\alpha \in \Phi$, then s_α leaves Φ invariant.

Proof. If α or φ is zero, then it is clear that $s_\alpha(\varphi)$ belongs to Φ. Thus we can suppose that $\alpha \in \Delta$ and proof that s_α leaves Δ invariant. For, we use the theorem 3.151 to find

$$s_\alpha\beta = \beta - \frac{2(\beta,\alpha)}{|\alpha|^2}\alpha = \beta - (\beta_\alpha - \beta^\alpha)\alpha. \tag{3.340}$$

If $\beta_\alpha - \beta^\alpha > 0$, we are in a case $\beta - n\alpha$ with $\beta_\alpha - \beta^\alpha < \beta_\alpha$, so that $s_\alpha\beta$ is a root. The case $\beta^\alpha > \beta_\alpha$ is treated in the same way. It just remains to check that if $\alpha,\beta \in \Delta$, then $s_\alpha\beta \neq 0$. The problem is to show that the equation (with a given α in Δ)

$$\beta = \frac{2(\alpha,\beta)}{(\alpha,\alpha)}\alpha \tag{3.341}$$

has no solution in Δ (the indeterminate is β). The only nonzero multiples of β which are roots are $\pm\beta$, then if we set $\beta = r\alpha$, equation (3.341) gives $r = \pm\frac{1}{2}$, which is impossible.

□

Proposition 3.155.
The Weyl group permutes simply transitively the simple systems.

3.9.4 Longest element

Let $w \in W$. The **length** of w is the smallest k such that w can be written as a composition of k reflexions s_{α_i}. That is the smallest k such that

$$w = s_{\alpha_{i_1}} s_{\alpha_{i_2}} \ldots s_{\alpha_{i_k}}. \tag{3.342}$$

Lemma 3.156.
If w and w' are elements of the Weyl group,

(i) $l(w) = l(w^{-1})$,

(ii) $l(w) = 0$ *if and only if* $w = \mathrm{id}$,

(iii) $l(ww') \leqslant l(w) + l(w')$,

(iv) $l(ww') \geqslant l(w) - l(w')$,

(v) $l(w) - 1 \leqslant l(ws_{\alpha_i}) \leqslant l(w) + 1$.

Let $n(w)$ be the number of positive simple roots that are send to a negative root:

$$n(w) = Card\,\Pi \cap w^{-1}(-\Pi). \tag{3.343}$$

Proposition 3.157.
Let Δ be a system of simple roots and Π the associated positive system. The following conditions on an element w of the Weyl group are equivalent:

(i) $w\Pi = \Pi$;

(ii) $w\Delta = \Delta$;

(iii) $l(w) = 0$;

(iv) $n(w) = 0$;

(v) $w = \mathrm{id}$.

For a proof see page 15 in [23].

Theorem 3.158.
If w is an element of the Weyl group,

$$l(w) = n(w). \tag{3.344}$$

Proof. No proof. □

3.9.5 Weyl group and representations

This subsection comes from [19].

Theorem 3.159.
There exists an irreducible representation of highest weight Λ if and only if

$$\Lambda_\alpha = \frac{2(\Lambda, \alpha)}{(\alpha, \alpha)} \in \mathbb{N} \tag{3.345}$$

for every simple root α. Moreover, if ξ is a highest weight vector and if α is a simple root, then

$$E_{-\alpha}^k \xi \begin{cases} \neq 0 & \text{if } k \leqslant \Lambda_\alpha \\ = 0 & \text{if } k > \Lambda_\alpha. \end{cases} \tag{3.346}$$

Proof. No proof. □

Theorem 3.160.
If Λ is the highest weight of a representation and if w_0 is the longest element of the Weyl group, then $w_0\Lambda$ is the lowest weight.

Problem and misunderstanding 15.
It is still not clear for me how does the proof works. Questions to be answered:

(i) existence, unicity

(ii) w_0 is the longest element of the Weyl group

(iii) if Λ is the highest weight, then $w_0\Lambda$ is the lowest.

3.9.6 Chevalley basis (deprecated)

See [24].

Let Φ be the finite set of roots of $\mathfrak{g}$. Then chose a positivity notion on $\mathfrak{h}^*$ and consider Φ^+, the positive subset of Φ. We also take Δ, a basis of the roots. An element of Φ^+ is a **simple root** if it cannot be written under the form of a sum of two elements of Φ^+. Every positive root is a sum of simple roots.

Let

$$\{\alpha_1,\ldots,\alpha_l\} \tag{3.347}$$

be a basis of $\mathfrak{h}^*$ made of simple roots and

$$\{h_1,\ldots,h_l\}, \tag{3.348}$$

the dual basis. One can choose the α_i in such a way that $\{h_1,\ldots,h_l\}$ is orthogonal with respect to the Killing form[23]. One consequence of that is that

$$B(h_i,h)=\alpha_i(h) \tag{3.349}$$

for every $h\in\mathfrak{h}$. Indeed, h can be written, in the basis, as $h=h^jh_j$ where $h^j=B(h_j,h)$. Thus one has

$$B(h_i,h)=h^i=h^j\delta_{ij}=\alpha_i(h^jh_j)=\alpha_i(h). \tag{3.350}$$

We consider $\{\alpha_1,\ldots,\alpha_m\}$, the positive roots (the roots α_1,…,α_l are some of them). One knows that $\mathfrak{g}_{\alpha_i}$ is one dimensional, so one take $e_i\in\mathfrak{g}_{\alpha_i}$ and $f_i\in\mathfrak{g}_{-\alpha_i}$ as basis of their respective spaces. If we denote by $\mathfrak{n}^+=\mathrm{Span}\{e_1,\ldots,e_m\}$ and $\mathfrak{n}^-=\mathrm{Span}\{f_1,\ldots,f_m\}$, we have the decomposition

$$\mathfrak{g}=\mathfrak{n}^-\oplus\mathfrak{h}\oplus\mathfrak{n}^+. \tag{3.351}$$

It $\{\alpha_i\}$ are the simple roots, we consider the following new basis for $\mathfrak{h}$:

$$H_{\alpha_i}=\frac{2\alpha_i^*}{(\alpha_i,\alpha_i)} \tag{3.352}$$

where α_i^* is the dual of α_i with respect to the inner product on $\mathfrak{h}^*$, this means

$$\alpha_j(\alpha_i^*)=(\alpha_i,\alpha_j). \tag{3.353}$$

Since $\mathfrak{h}$ is abelian (proposition 3.149), we have

$$[H_{\alpha_i},H_{\alpha_j}]=0. \tag{3.354}$$

Each root is a combination of the simple roots. If $\beta=\sum_{i=1}^l k_i\alpha_i$, we generalise the definition of H_{α_i} to

$$H_\beta=\frac{2\beta^*}{(\beta,\beta)}=\sum_i k_i\frac{(\alpha_i,\alpha_i)}{(\beta,\beta)}H_{\alpha_i}. \tag{3.355}$$

The element H_β is the **co-weight** associated with the weight β.

Using the inner product $(.,.)$, we have the decomposition $\beta=\sum_i(\beta,\alpha_i)\alpha_i$ of the roots. An immediate consequence is that

$$\beta(\alpha_i^*)=(\alpha_i,\beta). \tag{3.356}$$

If β is any root, we denote by β_i the result of β on H_{α_i}:

$$\beta_i=\beta(H_{\alpha_i})=\frac{2(\alpha_i,\beta)}{(\alpha_i,\alpha_i)}. \tag{3.357}$$

[23] Why ?

Theorem 3.161 (Chevalley basis).
For each root β, one can found an eigenvector E_β of $\mathrm{ad}(H_\beta)$ such that

$$\begin{aligned}
[H_\beta, H_\gamma] &= 0\\
[E_\beta, E_{-\beta}] &= H_\beta\\
[E_\beta, E_\gamma] &= \begin{cases} \pm(p+1)E_{\beta+\gamma} & \text{if } \beta+\gamma \text{ is a root}\\ 0 & \text{otherwise}\end{cases}\\
[H_\beta, E_\gamma] &= 2\frac{(\beta,\gamma)}{(\beta,\beta)}E_\gamma
\end{aligned} \tag{3.358}$$

where p is the biggest integer j such that $\gamma + j\beta$ is a root. Moreover, if α_i and α_j are simple roots, the latter becomes

$$[H_{\alpha_i}, E_{\pm\alpha_j}] = \pm A_{ij} E_{\pm\alpha_j} \tag{3.359}$$

where A is the Cartan matrix.

An important point to notice is that, for each positive root α, the algebra generated by $\{H_\alpha, E_\alpha, E_{-\alpha}\}$ is $\mathfrak{sl}(2)$. This is the reason why the representation theory of $\mathfrak{g}$ reduces to the representation theory of $\mathfrak{sl}(2)$.

3.10 Real Lie algebras

3.10.1 Real and complex vector spaces

If V is a real vector space, the **complexification** of V is the vector space

$$V^{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C}.$$

If $\{v_i\}$ is a basis of V on $\mathbb{R}$, then $\{v_i \otimes 1\}$ is a basis of $V^{\mathbb{C}}$ on C. Then

$$\dim_{\mathbb{R}} V = \dim_{\mathbb{C}} V^{\mathbb{C}}.$$

Let W be a complex vector space. If one restrains the scalars to $\mathbb{R}$, we find a real vector space denoted by $W^{\mathbb{R}}$. If $\{w_j\}$ is a basis of W, then $\{w_j, iw_j\}$ is a basis of $W^{\mathbb{R}}$ and

$$\dim_{\mathbb{C}} W = \frac{1}{2}\dim_{\mathbb{R}} W^{\mathbb{R}}.$$

Note that $(V^{\mathbb{C}})^{\mathbb{R}} = V \oplus iV$.

A real vector space V is a **real form** of a complex vector space W if $W^{\mathbb{R}} = V \oplus iV$. If V is a real form of W, the map $\varphi\colon V^{\mathbb{C}} \to V^{\mathbb{C}}$ given by the identity on V and the multiplication by -1 on iV is the **conjugation** of $V^{\mathbb{C}}$ with respect of the real form V.

3.10.2 Real and complex Lie algebras

For notational convenience, if not otherwise mentioned, $\mathfrak{g}$ will denote a complex Lie algebra and $\mathfrak{f}$ a real one. If $\mathfrak{f}$ is a real Lie algebra and $\mathfrak{f}^{\mathbb{C}} = \mathfrak{f} \otimes \mathbb{C}$, its complexification (as vector space), we endow $\mathfrak{f}^{\mathbb{C}}$ with a Lie algebra structure by defining

$$[(X \otimes a), (Y \otimes b)] = [X, Y] \otimes ab.$$

This is a bilinear extension of the Lie algebra bracket of $\mathfrak{f}$. It is rather easy to see that $[\mathfrak{f}, \mathfrak{f}]^{\mathbb{C}} = [\mathfrak{f}^{\mathbb{C}}, \mathfrak{f}^{\mathbb{C}}]$.

Now we turn our attention to the Killing form. Let $\mathfrak{f}$ be a real Lie algebra with a Killing form $B_{\mathfrak{f}}$. A basis of $\mathfrak{f}$ is also a basis of $\mathfrak{f}^{\mathbb{C}}$. Then the matrix $B_{ij} = \mathrm{Tr}(\mathrm{ad}\, X_i \circ \mathrm{ad}\, X_j)$ of the Killing form is the same for $\mathfrak{f}^{\mathbb{C}}$ than for $\mathfrak{f}$. In conclusion:

$$B_{\mathfrak{f}^{\mathbb{C}}}|_{\mathfrak{f}\times\mathfrak{f}} = B_{\mathfrak{f}}.$$

Let us study the inverse process: $\mathfrak{g}$ is a complex Lie algebra and $\mathfrak{g}^{\mathbb{R}}$ is the real Lie algebra obtained from $\mathfrak{g}$ by restriction of the scalars. If $\mathcal{B} = \{v_j\}$ is a basis of $\mathfrak{g}$, $\mathcal{B}' = \{v_j, iv_j\}$ is a one of $\mathfrak{g}^{\mathbb{R}}$. For a certain $X \in \mathfrak{g}$ we denote by (c_{kl}) the matrix of $\mathrm{ad}_{\mathfrak{g}} X$. Now we study the matrix of $\mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}} X$ in the basis $\mathcal{B}'$ by computing

$$(\mathrm{ad}_{\mathfrak{g}} X)v_i = c_{ik}v_k = \big[\mathrm{Re}(c_{ik}) + i\,\mathrm{Im}(c_{ik})\big]v_k = a_{ik}v_k + b_{ik}(iv_k) \tag{3.360}$$

if $a = \mathrm{Re}\, c$ and $b = \mathrm{Im}\, c$. Then the columns of $\mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}}$ which correspond to the $v_i \in \mathcal{B}'$'s are given by

$$\mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}} X = \begin{pmatrix} a & \cdot \\ b & \cdot \end{pmatrix}$$

where the dots denote some entries to be find now:

$$(\mathrm{ad}_{\mathfrak{g}} X)(iv_i) = i\big(a_{ik}v_k + b_{ik}(iv_k)\big) = a_{ik}(iv_k) - b_{ik}v_k, \tag{3.361}$$

so that the complete matrix of $\mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}} X$ in the basis $\mathcal{B}'$ is given by

$$\mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}} X = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

So,

$$\mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}} X \circ \mathrm{ad}_{\mathfrak{g}^{\mathbb{R}}} X' = \begin{pmatrix} aa' - bb' & \cdot \\ \cdot & aa' - bb' \end{pmatrix}.$$

Then $B(X, X') = 2\,\mathrm{Tr}(aa' - bb')$ while

$$B(X, Y) = \mathrm{Tr}\,\big((a + ib)(a' + ib')\big) = \mathrm{Tr}(aa' - bb') + i\,\mathrm{Tr}(ab' + ba'). \tag{3.362}$$

Thus we have

$$B_{\mathfrak{g}^{\mathbb{R}}} = 2\,\mathrm{Re}\,B_{\mathfrak{g}}, \tag{3.363}$$

so that $\mathfrak{g}^{\mathbb{R}}$ is semisimple if and only if $\mathfrak{g}$ is semisimple.

A result about the group of inner automorphism which will be useful later:

Lemma 3.162.
If $\mathfrak{g}$ is a complex semisimple Lie algebra, then $\mathrm{Int}\,\mathfrak{g} = \mathrm{Int}\,\mathfrak{g}^{\mathbb{R}}$.

Proof. If $\{X_i\}$ is a basis of $\mathfrak{g}$, then $\{X_j, iX_j\}$ is a basis of $\mathfrak{g}^{\mathbb{R}}$. We define $\psi\colon \mathrm{ad}\,\mathfrak{g} \to \mathrm{ad}\,\mathfrak{g}^{\mathbb{R}}$ by

$$\psi(\mathrm{ad}(a^j X_j)) = \mathrm{ad}(a^j X_j).$$

It is clearly surjective. On the other hand, if $\mathrm{ad}(a^j X_j)\,\mathrm{ad}(b^k X_k)$ as elements of $\mathrm{ad}\,\mathfrak{g}^{\mathbb{R}}$, then they are equals as elements of $\mathrm{ad}\,\mathfrak{g}$. The discussion following equations (3.6) finish the proof. □

3.10.3 Split real form

Let $\mathfrak{g}$ be a complex semisimple Lie algebra, $\mathfrak{h}$ a Cartan subalgebra, Φ the set roots, Δ the set of non zero roots and B, the Killing form. From property (3.300) and the fact that $c(-\alpha, -\beta) = c(\alpha, \beta)$, we find $c(\alpha, \beta)^2 = \frac{1}{2}\beta^\alpha(1 + \beta_\alpha)|\alpha|^2$, so that $c(\alpha, \beta)^2 \geqslant 0$ which gives $c(\alpha, \beta) \in \mathbb{R}$. We can define

$$\mathfrak{g}_{\mathbb{R}} = \mathfrak{h}_0 \bigoplus_{\alpha \in \Phi} \mathbb{R}x_\alpha.$$

Remark that $\mathfrak{g}_\alpha$ has dimension one with respect to $\mathbb{C}$, not $\mathbb{R}$; then $\mathbb{R}x_\alpha \neq \mathfrak{g}_\alpha$, but $\mathbb{C}x_\alpha = \mathfrak{g}_\alpha$ and $\mathfrak{g}_\alpha = \mathbb{R}x_\alpha \oplus i\mathbb{R}x_\alpha$. Since it is clear that $\bigoplus_{\alpha \in \Delta}(\mathbb{R}x_\alpha \oplus i\mathbb{R}x_\alpha) = \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha$, the proposition 3.145 gives

$$\mathfrak{g} = \mathfrak{g}_{\mathbb{R}} \oplus i\mathfrak{g}_{\mathbb{R}}. \tag{3.364}$$

Any real form of $\mathfrak{g}$ which contains the $\mathfrak{h}_{\mathbb{R}}$ of a certain Cartan subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ is said a **split real form**. The construction shows that any complex semisimple Lie algebra admits a split real form.

3.10.4 Compact real form

A **compact real form** of a complex Lie algebra is a real form which is compact as Lie algebra. Recall that a real Lie algebra is compact when its analytic group of inner automorphism is compact, see page 56

Theorem 3.163.
Any complex semisimple Lie algebra contains a compact real form.

Proof. Let $\mathfrak{h}$ be a Cartan algebra of the complex semisimple Lie algebra $\mathfrak{g}$ and x_α, some root vectors. We consider the space

$$\mathfrak{u}_0 = \underbrace{\sum_{\alpha \in \Phi} \mathbb{R}ih_\alpha}_{A} + \underbrace{\sum_{\alpha \in \Phi} \mathbb{R}(x_\alpha - x_{-\alpha})}_{B} + \underbrace{\sum_{\alpha \in \Phi} \mathbb{R}i(x_\alpha + x_{-\alpha})}_{C}. \tag{3.365}$$

Since $\mathfrak{u}_0 \oplus i\mathfrak{u}_0$ contains all the $\mathbb{C}h_\alpha$, $\mathfrak{h} \subset \mathfrak{u}_0 \oplus i\mathfrak{u}_0$; it is also rather clear that $\mathfrak{u}_0$ is a real form of $\mathfrak{g}$ (as vector space), for example, $i\mathbb{R}(x_\alpha - x_{-\alpha}) + \mathbb{R}i(x_\alpha + x_{-\alpha}) = \mathbb{R}ix_\alpha$. Now we have to check that $\mathfrak{u}_0$ is a real form of $\mathfrak{g}$ as Lie algebra, i.e. that $\mathfrak{u}_0$ is closed for the Lie bracket. This is a lot of computations:

$$\begin{aligned}
[ih_\alpha, ih_\beta] &= 0,\\
[ih_\alpha, (x_\alpha - x_{-\alpha})] &= i(\alpha(h_\alpha)x_\alpha - (-\alpha)(h_\alpha)x_{-\alpha})\\
&= i\alpha(h_\alpha)(x_\alpha + x_{-\alpha}) \in C,\\
[ih_\alpha, i(x_\alpha + x_{-\alpha})] &= -\alpha(h_\alpha)(x_\alpha - x_{-\alpha}) \in B,\\
[(x_\alpha - x_{-\alpha}), (x_\beta - x_{-\beta})] &= c(\alpha,\beta)(x_{\alpha+\beta} - x_{-(\alpha+\beta)}) \in B\\
&\quad - c(\alpha,\beta)(x_{\alpha-\beta} - x_{\beta-\alpha}) \in B,\\
[(x_\alpha - x_{-\alpha}), i(x_\beta + x_{-\beta})] &= ic(\alpha,\beta)(x_{\alpha+\beta} + x_{-(\alpha+\beta)}) \in C\\
&\quad + ic(\alpha,-\beta)(x_{\alpha-\beta)} + x_{-\alpha+\beta}) \in C\\
[ih_\alpha, (x_\beta - x_{-\beta})] &= i\beta(h_\alpha)(x_\beta - x_{-\beta}) \in C\\
[ih_\alpha, i(x_\beta + x_{-\beta})] &= -\beta(h_\alpha)(x_\beta - x_{-\beta}) \in B\\
[i(x_\alpha + x_{-\alpha}), i(x_\beta + x_{-\beta})] &= -c(\alpha,\beta)(x_{\alpha+\beta} - x_{-(\alpha+\beta)})\\
&\quad - c'(\alpha,-\beta)(x_{\alpha-\beta} - x_{-\alpha+\beta}).
\end{aligned}$$

From proposition 3.64, it just remains to prove that the Killing form of $\mathfrak{u}_0$ is strictly negative definite. We know that $B_\mathfrak{g}(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ if $\alpha, \beta \in \Phi$ and $\alpha + \beta \neq 0$; then $A \perp B$ and $A \perp C$. It is a lot of computation to compute the Killing form; we know that B is strictly positive definite on $\sum_{\alpha\in\Delta} \mathbb{R}h_\alpha$ (and then strictly negative definite on A) a part this, the non zero elements are (recall that if $\alpha \neq 0$, $B(x_\alpha, x_\alpha) = 0$ from corollary 3.79)

$$\begin{aligned}
&B((x_\alpha - x_{-\alpha}), (x_\alpha - x_{-\alpha})) = -2B(x_\alpha, x_{-\alpha}) = -2\\
&B(i(x_\alpha + x_{-\alpha}), i(x_\alpha, x_{-\alpha})) = -2.
\end{aligned}$$

What we have in the matrix of $B_\mathfrak{g}|_{\mathfrak{u}_0\times\mathfrak{u}_0}$ is a negative definite block (corresponding to A), -2 on the rest of the diagonal and zero anywhere else. Then it is well negative definite and $\mathfrak{u}_0$ is a compact real from of $\mathfrak{g}$. □

3.10.5 Involutions

Let $\mathfrak{g}$ be a (real or complex) Lie algebra. An automorphism $\sigma\colon \mathfrak{g} \to \mathfrak{g}$ which is not the identity such that σ^2 is the identity is a **involution**. An involution $\theta\colon \mathfrak{k} \to \mathfrak{k}$ of a *real* semisimple Lie algebra $\mathfrak{k}$ such that the quadratic form B_θ defined by

$$B_\theta(X,Y) := -B(X, \theta Y)$$

is positive definite is a **Cartan involution**.

Proposition 3.164.
Let $\mathfrak{g}$ be a complex semisimple Lie algebra, $\mathfrak{u}_0$ a compact real form and τ, the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{u}_0$. Then τ is a Cartan involution of $\mathfrak{g}^\mathbb{R}$.

Proof. From the assumptions, $\mathfrak{g} = \mathfrak{u}_0 \oplus i\mathfrak{u}_0$, $\tau_{\mathfrak{u}_0} = id$ and $\tau_{i\mathfrak{u}_0} = -id$; then it is clear that $\tau^2_{\mathfrak{g}^\mathbb{R}} = id|_{\mathfrak{g}^\mathbb{R}}$. If $Z \in \mathfrak{g}$, we can decompose into $Z = X + iY$ with X, $Y \in \mathfrak{u}_0$. For $Z \neq 0$, we have

$$B_\mathfrak{g}(Z, \tau Z) = B_\mathfrak{g}(X + iY, X - iY) = B_\mathfrak{g}(X,X) + B_\mathfrak{g}(Y,Y) = B_{\mathfrak{u}_0}(X,X) + B_{\mathfrak{u}_0}(Y,Y) < 0 \tag{3.366}$$

because B restricts itself to $\mathfrak{u}_0$ which is compact. Then

$$(B_{\mathfrak{g}^\mathbb{R}})_\tau(Z, Z') = B_{\mathfrak{g}^\mathbb{R}}(Z, \tau Z) = -2\operatorname{Re} B_\mathfrak{g}(Z, \tau Z') \tag{3.367}$$

is positive definite because $(B_\mathfrak{g})_\tau$ is negative definite. Thus τ is a Cartan involution of $\mathfrak{g}^\mathbb{R}$. □

Lemma 3.165.
If φ and ψ are involutions of a vector space V (we denote by V_{ψ^+} and V_{ψ^-} the subspaces of V for the eigenvalues 1 and -1 of ψ and similarly for φ), then

$$[\varphi, \psi] = 0 \quad \textit{iff} \quad \begin{cases} V_{\varphi^+} = (V_{\varphi^+} \cap V_{\psi^+}) \oplus (V_{\varphi^+} \cap V_{\psi^-})\\ V_{\varphi^-} = (V_{\varphi^-} \cap V_{\psi^+}) \oplus (V_{\varphi^-} \cap V_{\psi^-}), \end{cases}$$

i.e. if and only if the decomposition of V with respect to φ is "compatible" with the one with respect to ψ.

Proof. Direct sense. Let us first see that φ leaves the decomposition $V = V_{\psi^+} \oplus V_{\psi^-}$ invariant. If $x = x_{\psi^+} + x_{\psi^-}$,

$$\varphi(x_{\psi^+}) = (\varphi \circ \psi)(x_{\psi^+}) = (\psi \circ \varphi)(x_{\psi^+}).$$

Then $\varphi(x_{\psi^+}) \in V_{\psi^+}$, and the matrix of φ is block-diagonal with respect to the decomposition given by ψ. Thus V_{ψ^+} and V_{ψ^-} split separately into two parts with respect to φ.
Inverse sense. If $x \in V$, we can write $x = x_{++} + x_{+-} + x_{-+} + x_{--}$ where the first index refers to ψ while the second one refers to ψ; for example, $x_{+-} \in V_{\psi^+} \cap V_{\varphi^-}$. The following computation is easy:

$$\begin{aligned}(\varphi \circ \psi)(x) &= \varphi(x_{++} + x_{+-} - x_{-+} - x_{--}) \\ &= x_{++} - x_{+-} - x_{-+} + x_{--} \\ &= \psi(x_{++} - x_{+-} - x_{-+} - x_{--}) \\ &= (\psi \circ \varphi)(x).\end{aligned} \tag{3.368}$$

□

Theorem 3.166.
Let $\mathfrak{f}$ be a real semisimple Lie algebra, θ a Cartan involution on $\mathfrak{f}$ and σ, another involution (not specially Cartan). Then there exists a $\varphi \in \operatorname{Int}\mathfrak{f}$ such that $[\varphi\theta\varphi^{-1}, \sigma] = 0$

Proof. If θ is a Cartan involution, then B_θ is a scalar product on $\mathfrak{f}$. Let $\omega = \sigma\theta$. By using $\sigma^2 = \theta^2 = 1$, $\theta = \theta^{-1}$ and the invariance property 3.12 of the Killing form,

$$B(\omega X, \theta Y) = B(X, \omega^{-1}\theta Y) = B(X, \theta\sigma\theta Y) = B(X, \theta\omega Y). \tag{3.369}$$

Then $B_\theta(\omega X, Y) = B_\theta(X, \omega Y)$. This is a general property of scalar product that in this case, the matrix of ω is symmetric while the one of ω^2 is positive definite. If we consider the classical scalar product whose matrix is (δ_{ij}), the property is written as $A_{ij}v_j w_j = v_i A_{ij} w_j$ (with sum over i and j); this implies the symmetry of A. To see that A^2 is positive definite, we compute (using the symmetry):

$$A_{ij}A_{jk}v_i v_k = v_i A_{ij} v_k A_{kj} = \sum_j (v_i A_{ij})^2 > 0.$$

The next step is to see that there is an unique linear transformation $A\colon \mathfrak{f} \to \mathfrak{f}$ such that $\omega^2 = e^A$, and that for any $t \in \mathbb{R}$, the transformation e^{tA} is an automorphism of $\mathfrak{f}$.

We choose an orthonormal (with respect to the inner product B_θ) basis $\{X_1, \ldots, X_n\}$ of $\mathfrak{f}$ in which ω is diagonal. In this basis, ω^2 is also diagonal and has positive real numbers on the diagonal; then the existence and unicity of A is clear. Now we take some notations:

$$\omega(X_i) = \lambda_i X_i \tag{3.370a}$$
$$\omega^2(X_i) = e^{a_i} X_i, \tag{3.370b}$$

(no sum at all) where the a_i are the diagonals elements of A. The structure constants are as usual defined by

$$[X_i, X_j] = c_{ij}^k X_k. \tag{3.371}$$

Since σ and θ are automorphisms, ω^2 is also one. Then

$$\omega^2[X_i, X_j] = c_{ij}^k \omega^2(X_k) = c_{ij}^k e^{a_k} X_k$$

can also be computed as

$$\omega^2[X_i, X_j] = [\omega^2 X_i, \omega^2 X_j] = e^{a_i} e^{a_j} c_{ij}^k X_k,$$

so that $c_{ij}^k e^{a_k} = c_{ij}^k e^{a_i} e^{a_j}$, and then $\forall t \in \mathbb{R}$,

$$c_{ij}^k e^{t a_k} = c_{ij}^k e^{t a_i} e^{t a_j},$$

which proves that e^{tA} is an automorphism of $\mathfrak{f}$. By lemma 2.17, A is thus a derivation of $\mathfrak{f}$. The semi-simplicity makes $\partial\mathfrak{f} = \operatorname{ad}\mathfrak{f}$, then $A \in \operatorname{ad}\mathfrak{f}$ and $e^{tA} \in \operatorname{Int}\mathfrak{f}$ because it clearly belongs to the identity component of $\operatorname{Aut}\mathfrak{f}$.

Now we can finish de proof by some computations. Remark that $\omega = e^{A/2}$ and $[e^{tA}, \omega] = 0$ because it can be seen as a common matrix commutator. Since $\omega^{-1} = \theta\sigma$, we have $\theta\omega^{-1}\theta = \sigma\theta$, or $\theta\omega^2\theta = \omega^2$ and

$$e^A\theta = \theta e^{-A}. \tag{3.372}$$

From this, one can deduce that $e^{tA}\theta = \theta e^{-tA}$. Indeed, as matricial identity, equation (3.372) reads

$$(e^A\theta)_{ik} = (e^A)_{ij}\theta_{jk} = e^{a_i}\theta_{ik} = e^{-a_k}\theta_{ik}.$$

Then for any ik such that $\theta_{ik} \neq 0$, we find $e^{a_i} = e^{-a_k}$ and then also $e^{ta_i} = e^{-ta_k}$. Thus $(e^{tA}\theta)_{ik} = (e^{tA})_i j\theta_{jk} = e^{ta_i}\theta_{ik} = \theta_{ik}e^{-ta_k} = (\theta e^{-tA})_{ik}$. So we have

$$e^{tA}\theta = \theta e^{-tA}. \tag{3.373}$$

Now we consider $\varphi = e^{A/4} \in \operatorname{Int}\mathfrak{f}$ and $\theta_1 = \varphi\theta\varphi^{-1}$. We find $\theta_1\sigma = e^{A/2}\omega^{-1}$ and $\sigma\theta^{-1} = e^{-A/2}\omega$. Since $\omega^2 = A$, we have $e^{A/2} = e^{-A/2}\omega^2$ and thus $\theta_1\sigma = \sigma\theta_1$.

□

Corollary 3.167.
Any real Lie algebra has a Cartan involution.

Proof. Let $\mathfrak{f}$ be a real Lie algebra and $\mathfrak{g}$ be his complexification: $\mathfrak{g} = \mathfrak{f}^{\mathbb{C}}$. Let $\mathfrak{u}_0$ be a compact real form of $\mathfrak{g}$ and τ the induced involution (the conjugation) on $\mathfrak{g}$. By the proposition 3.164, we know that τ is a Cartan involution of $\mathfrak{g}^{\mathbb{R}}$. We also consider σ, the involution of $\mathfrak{g}$ with respect to the real form $\mathfrak{f}$. It is in particular an involution on the real Lie algebra $\mathfrak{f}$. Then one can find a $\varphi \in \operatorname{Int}\mathfrak{g}^{\mathbb{R}}$ such that $[\varphi\tau\varphi^{-1}, \sigma] = 0$ on $\mathfrak{g}^{\mathbb{R}}$. Let $\mathfrak{u}_1 = \varphi\mathfrak{u}_0$ and $X \in \mathfrak{u}_1$. We can write $X = \varphi Y$ for a certain $Y \in \mathfrak{u}_0$. Then

$$\varphi\tau\varphi^{-1}X = \varphi\tau Y = \varphi Y = X,$$

so that $\varphi\tau\varphi^{-1} = id|_{\mathfrak{u}_1}$. Note that $\mathfrak{u}_1$ is also a real compact form of $\mathfrak{g}$ because the Killing form is not affected by φ. Let τ_1 be the involution of $\mathfrak{g}$ induced by $\mathfrak{u}_1$. We have

$$\tau_1|_{\mathfrak{u}_1} = \varphi\tau\varphi^{-1}_{\mathfrak{u}_1} = \operatorname{id}|_{\mathfrak{u}_1}.$$

Since φ is $\mathbb{C}$-linear, we have in fact $\tau_1 = \varphi\tau\varphi^{-1}$. Now we forget $\mathfrak{u}_0$ and we consider the compact real form $\mathfrak{u}_1$ with his involution τ_1 of $\mathfrak{g}$ which satisfy $[\tau_1, \sigma] = 0$ on $\mathfrak{g}^{\mathbb{R}}$ This relation holds also on $i\mathfrak{g}^{\mathbb{R}}$, then

$$[\tau_1, \sigma] = 0$$

on $\mathfrak{g} = \mathfrak{f}^{\mathbb{C}}$. Let $X \in \mathfrak{f}$, i.e. $\sigma X = X$; it automatically fulfils

$$\sigma\tau_1 X = \tau_1\sigma X = \tau_1 X,$$

so that τ_1 restrains to an involution on $\mathfrak{f}$ (because $\tau_1\mathfrak{f} \subset \mathfrak{f}$). Let $\theta = \tau_1|_{\mathfrak{f}}$. For $X, Y \in \mathfrak{f}$, we have

$$B_\theta(X,Y) = -B_{\mathfrak{f}}(X,\theta Y) = -B_{\mathfrak{f}}(X,\tau Y) = \frac{1}{2}(B_{\mathfrak{g}^{\mathbb{R}}})_{\tau_1}(X,Y), \tag{3.374}$$

which shows that θ is a Cartan involution. The half factor on the last line comes from the fact that $\mathfrak{g}^{\mathbb{R}} = (\mathfrak{f}^{\mathbb{C}})^{\mathbb{R}} = \mathfrak{f} \oplus i\mathfrak{f}$.

□

Corollary 3.168.
Any two Cartan involutions of a real semisimple Lie algebra are conjugate by an inner automorphism.

Proof. Let σ and σ' be two Cartan involutions of $\mathfrak{f}$. We can find a $\varphi \in \inf\mathfrak{f}$ such that $[\varphi\sigma\varphi^{-1}, \sigma'] = 0$. Thus it is sufficient to prove that any two Cartan involutions which commute are equals. So let us consider θ and θ', two Cartan involutions such that $[\theta, \theta'] = 0$. By lemma 3.165, we know that the decompositions into $+1$ and -1 eigenspaces with respect to θ and θ' are compatibles. If we consider $X \in \mathfrak{f}$ such that $\theta X = X$ and $\theta' X = -1$ (it is always possible if $\theta \neq \theta'$), we have

$$\begin{aligned}0 < B_\theta(X,X) = -B(X,\theta X) = -B(X,X)\\0 < B_{\theta'}(X,X) = -B(X,\theta' X) = B(X,X)\end{aligned}$$

which is impossible. □

Corollary 3.169.
Any two real compact form of a complex semisimple Lie algebra are conjugate by an inner automorphism.

Proof. We know that any real form of $\mathfrak{g}$ induces an involution (the conjugation) and that if the real form is compact, the involution is Cartan on $\mathfrak{g}^{\mathbb{R}}$. Let $\mathfrak{u}_0$ and $\mathfrak{u}_1$ be two compact real forms of $\mathfrak{g}$ and τ_0, τ_1 the associated involutions of $\mathfrak{g}$ (which are Cartan involutions of $\mathfrak{g}^{\mathbb{R}}$). For a suitable $\varphi \in \operatorname{Int}\mathfrak{g}^{\mathbb{R}}$,

$$\tau_0 = \varphi\tau_1\varphi^{-1}.$$

The fact that $\operatorname{Int}\mathfrak{g} = \operatorname{Int}\mathfrak{g}^{\mathbb{R}}$ (lemma 3.162) finish the proof.

□

3.10.6 Cartan decomposition

Examples of Cartan and Iwasawa decomposition are given in sections ??, ??,??, ?? and ??. An example of how it works to prove isomorphism of Lie algebras is provided in subsection ??.

Let $\mathfrak{f}$ be a real semisimple Lie algebra. A vector space decomposition $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$ is a **Cartan decomposition** if the Killing form is negative definite on $\mathfrak{k}$ and positive definite on $\mathfrak{p}$ and the following commutators hold:

$$[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subseteq \mathfrak{p}, \quad [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}. \tag{3.375}$$

If $X \in \mathfrak{k}$ and $Y \in \mathfrak{p}$, we have $(\operatorname{ad} X \circ \operatorname{ad} Y)\mathfrak{k} \subseteq \mathfrak{p}$ and $(\operatorname{ad} X \circ \operatorname{ad} Y)\mathfrak{p} \subseteq \mathfrak{k}$, therefore $B_{\mathfrak{f}}(X, Y) = 0$.

Let $\theta\colon \mathfrak{f} \to \mathfrak{f}$ be a Cartan involution, $\mathfrak{k}$ its $+1$ eigenspace and $\mathfrak{p}$ his -1 one. It is easy to see that the relations (3.375) are satisfied for the decomposition $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$. For example, for $X, X' \in \mathfrak{k}$, using the fact that θ is an automorphism,

$$[X, X'] = [\theta X, \theta X'] = \theta[X, X'],$$

which proves that $[\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}$. Since θ is a Cartan involution, B_θ is positive definite. For $X \in \mathfrak{k}$,

$$B(X, X) = B(X, \theta X) = -B_\theta(X, X)$$

proves that B is negative definite on $\mathfrak{k}$; in the same way we find that B is also positive definite on $\mathfrak{p}$. Then the Cartan involution gives rise to a Cartan decomposition. We are going to prove that any Cartan decomposition defines a Cartan involution.

Let us now do the converse. Let $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$ be a Cartan decomposition of the real semisimple Lie algebra $\mathfrak{f}$. We define $\theta = \operatorname{id}|_{\mathfrak{k}} \oplus (-\operatorname{id})|_{\mathfrak{p}}$. If $X, X' \in \mathfrak{k}$, the definition of a Cartan algebra makes $[X, X'] \in \mathfrak{k}$ and so

$$\theta[X, X'] = [X, X'] = [\theta X, \theta X'],$$

and so on, we prove that θ is an automorphism of $\mathcal{F}$. It remains to prove that B_θ is positive definite. If $X \in \mathfrak{k}$,

$$B_\theta(X, X) = -B(X, \theta X) = -B(X, X).$$

Then B_θ is positive definite on $\mathfrak{k}$ because on this space, B is negative definite by definition of a Cartan involution. The same trick shows that B_θ is also positive definite on $\mathfrak{p}$. We had seen that $\mathfrak{p}$ and $\mathfrak{k}$ where B_θ-orthogonal spaces. Thus B_θ is positive definite and θ is a Cartan involution.

Let $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$ be a Cartan decomposition. Then it is quite easy to see that $\mathfrak{k} \oplus i\mathfrak{p}$ is a compact real form of $\mathfrak{g} = (\mathfrak{f}^{\mathbb{C}})$.

Proposition 3.170.
Let $\mathfrak{L}$ and $\mathfrak{q}$ be the $+1$ and -1 eigenspaces of an involution σ. Then σ is a Cartan involution if and only if $\mathfrak{L} \oplus i\mathfrak{q}$ is a compact real form of $\mathfrak{f}^{\mathbb{C}}$.

Proof. First remark that $\mathfrak{L} \oplus i\mathfrak{q}$ is always a real form of $\mathfrak{f}^{\mathbb{C}}$. The direct sense is yet done. Then we suppose that $B_{\mathfrak{f}^{\mathbb{C}}}$ is negative definite on $\mathfrak{L} \oplus i\mathfrak{q}$ and we have to show that $\mathfrak{L} \oplus \mathfrak{q}$ is a Cartan decomposition of $\mathfrak{f}$. The condition about the brackets on $\mathfrak{L}$ and $\mathfrak{q}$ is clear from their definitions. If $X \in \mathfrak{L}$, $B(X, X) < 0$ because B is negative definite on $\mathfrak{L}$. If $Y \in \mathfrak{q}$, $B(Y, Y) = -B(iY, iY) > 0$ because B is negative definite on $i\mathfrak{q}$. □

3.11 Root spaces in the real case

Let $\mathfrak{f}$ be a real semisimple Lie algebra with a Cartan involution θ and the corresponding Cartan decomposition $\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{p}$. We consider B, a "Killing like" form, i.e. B is a symmetric nondegenerate invariant bilinear form on $\mathfrak{f}$ such that $B(X, Y) = B(\theta X, \theta Y)$ and $B_\theta := -B(X, \theta X)$ is positive definite. Then B is negative definite on the compact real form $\mathfrak{k} \oplus i\mathfrak{p}$. Indeed if $Y \in \mathfrak{p}$,

$$B(iY, iY) = -B(\theta Y, \theta Y) = B(Y, \theta Y) = -B_\theta(Y, Y) < 0. \tag{3.376}$$

The case with $X \in \mathfrak{k}$ is similar. It is easy to see that B_θ is in fact a scalar product on $\mathfrak{f}$, so that we can define the orthogonality and the adjoint from B_θ. If $A\colon \mathfrak{f} \to \mathfrak{f}$ is an operator on $\mathfrak{f}$, his adjoint is the operator A^* given by the formula

$$B_\theta(AX, Y) = B_\theta(X, A^*Y)$$

for all X, $Y \in \mathfrak{f}$.

Proposition 3.171.
With this definition, when $X \in \mathfrak{f}$, the adjoint operator of $\operatorname{ad} X$ is given by means of the Cartan involution:

$$(\operatorname{ad} X)^* = \operatorname{ad}(\theta X).$$

Proof. This is a simple computation

$$B_\theta\big((\operatorname{ad}\theta X)Y, Z\big) = -B\big(Y, [\theta X, \theta Y]\big) = -B_\theta(Y, [X, Z]) = -B_\theta\big((\operatorname{ad} X)^* Y, Z\big). \tag{3.377}$$

□

Let $\mathfrak{a}$ be a maximal abelian subalgebra of $\mathfrak{p}$ (the existence comes from the finiteness of the dimensions). If $H \in \mathfrak{a}$, the operator $\operatorname{ad} H$ is self-adjoint because

$$(\operatorname{ad} H)^* X = (-\operatorname{ad}\theta H)X = [H, X] = (\operatorname{ad} H)X, \tag{3.378}$$

where we used the fact that $H \in \mathfrak{p}$. For $\lambda \in \mathfrak{a}^*$, we define the space

$$\mathfrak{f}_\lambda = \{X \in \mathfrak{f} \text{ st } \forall H \in \mathfrak{a}, (\operatorname{ad} H)X = \lambda(H)X\}. \tag{3.379}$$

If $\mathfrak{f}_\lambda \neq 0$ and $\lambda \neq 0$, we say that λ is a **restricted root** of $\mathfrak{f}$. We denote by Σ the set of restricted roots of $\mathfrak{f}$. We may sometimes write $\Sigma_\mathfrak{f}$ if the Lie algebra is ambiguous.

The main properties of the real root spaces are given in the following proposition.

Proposition 3.172.
The set Σ of the restricted roots of a real semisimple Lie algebra $\mathfrak{f}$ has the following properties:

(i) $\mathfrak{f} = \mathfrak{f}_0 \bigoplus_{\lambda\in\Sigma} \mathfrak{f}_\lambda$,

(ii) $[\mathfrak{f}_\lambda, \mathfrak{f}_\mu] \subseteq \mathfrak{f}_{\lambda+\mu}$,

(iii) $\theta\mathfrak{f}_\lambda = \mathfrak{f}_{-\lambda}$,

(iv) $\lambda \in \Sigma$ *implies* $-\lambda \in \Sigma$,

(v) $\mathfrak{f}_0 = \mathfrak{a} \oplus \mathfrak{m}$ *where* $\mathfrak{m} = \mathcal{Z}_\mathfrak{k}(\mathfrak{a})$ *and* $\mathfrak{a} \perp \mathfrak{m}$.

Proof. Proof of (i). The operators $\operatorname{ad} H$ with $H \in \mathfrak{a}$ form an abelian algebra of self-adjoint operators, then they are simultaneously diagonalisable. Let $\{X_i\}$ be a basis which realize this diagonalisation, and $\mathfrak{f}_i = \operatorname{Span} X_i$, so that $\mathfrak{f} = \oplus_i \mathfrak{f}_i$. We have $(\operatorname{ad} H)\mathfrak{f}_i = \mathfrak{f}_i$ and then $(\operatorname{ad} H)X_i = \lambda_i(H)X_i$ for a certain $\lambda_i \in \mathfrak{a}^*$. This shows that $\mathfrak{f}_i \subseteq \mathfrak{f}_{\lambda_i}$.[24]

Proof of (ii). Let $H \in \mathfrak{a}$, $X \in \mathfrak{f}_\lambda$ and $Y \in \mathfrak{f}_\mu$. We have

$$(\operatorname{ad} H)[X, Y] = [[H, X], Y] + [X, [H, Y]] = \big(\lambda(H) + \mu(H)\big)[X, Y]. \tag{3.380}$$

Proof of (iii). Using the fact that $\theta H = -H$ because $H \in \mathfrak{p}$,

$$(\operatorname{ad} H)\theta X = \theta[\theta H, X] = -\theta\lambda(H)X = -\lambda(H)(\theta X). \tag{3.381}$$

Proof of (iv). It is a consequence of (iii) because if $\mathfrak{f}_\lambda \neq 0$, then $\theta\mathfrak{f}_\lambda \neq 0$.

Proof of (v). By (iii), $\theta\mathfrak{f}_0 = \mathfrak{f}_0$, then $\mathfrak{f}_0 = (\mathfrak{k} \cap \mathfrak{f}_0) \oplus (\mathfrak{p} \cap \mathfrak{f}_0)$. If $X \in \mathfrak{f}_0$, then it commutes with all the elements of $\mathfrak{a}$ and by the maximality property of $\mathfrak{a}$, provided that $X \in \mathfrak{p}$, it also must belongs to $\mathfrak{a}$. This fact makes $\mathfrak{a} = \mathfrak{p} \cap \mathfrak{f}_0$. Now,

$$\mathfrak{m} = \mathcal{Z}_\mathfrak{k}(\mathfrak{a}) = \{X \in \mathfrak{k} \text{ st } [X, \mathfrak{a}] = 0\} = \mathfrak{k} \cap \mathfrak{f}_0.$$

All this gives $\mathfrak{f}_0 = \mathcal{Z}_\mathfrak{k}(\mathfrak{a}) \oplus \mathfrak{a}$. □

We choose a positivity notion on $\mathfrak{a}^*$, we consider Σ^+, the set of restricted positive roots and we define

$$\mathfrak{n} = \bigoplus_{\lambda\in\Sigma^+} \mathfrak{f}_\lambda.$$

From finiteness of the dimension, there are only a finitely many forms $\lambda \in \mathfrak{a}^*$ such that $\mathfrak{f}_\lambda \neq 0$. Then, taking, more and more commutators in $\mathfrak{n}$, the formula $[\mathfrak{f}_\lambda, \mathfrak{f}_\mu] \subseteq \mathfrak{f}_{\lambda+\mu}$ shows that the result finish to fall into a $\mathfrak{f}_\mu = 0$. On the other hand, since $\mathfrak{a} \subset \mathfrak{f}_0$, we have $[\mathfrak{a}, \mathfrak{n}] = \mathfrak{n}$. If $a_1, a_2 \in \mathfrak{a}$ and $n_1, n_2 \in \mathfrak{n}$,

$$[a_1 + n_1, a_2 + n_2] = \underbrace{[a_1, a_2]}_{=0} + \underbrace{[a_1, n_2]}_{\in\mathfrak{n}} + \underbrace{[n_1, a_2]}_{\in\mathfrak{n}} + \underbrace{[n_1, n_2]}_{\in\mathfrak{n}}, \tag{3.382}$$

then $[\mathfrak{a} \oplus \mathfrak{n}, \mathfrak{a} \oplus \mathfrak{n}] = \mathfrak{n}$. This proves the three following important properties:

(i) $\mathfrak{n}$ is nilpotent.

(ii) $\mathfrak{a}$ is abelian.

(iii) $\mathfrak{a} \oplus \mathfrak{n}$ is a solvable Lie subalgebra of $\mathfrak{f}$.

[24] pourquoi ça n'implique pas que $\dim \mathfrak{f}_{\lambda_i} = 1$? Réponse par Philippe : tu as oublié les valeurs propres nulles dans ta base ce qui doit entrainer quelques modifs dans ton texte(par ex. $adHf_i = f_i$ pas toujours)

3.11.1 Iwasawa decomposition

Theorem 3.173.
Let $\mathfrak{f}$ be a real semisimple Lie algebra and $\mathfrak{k}$, $\mathfrak{a}$, $\mathfrak{n}$ as before. Then we have the following direct sum:

$$\mathfrak{f} = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}. \tag{3.383}$$

This is the **Iwasawa decomposition** for the real semisimple Lie algebra $\mathfrak{f}$.

Proof. We yet know the direct sum $\mathfrak{f} = \mathfrak{f}_0 \bigoplus_{\lambda\in\Sigma} \mathfrak{f}_\lambda$. Roughly speaking, in $\mathfrak{n}$ we have only vectors of Σ^+, in $\theta\mathfrak{n}$, only of Σ^- and in $\mathfrak{a}$, only in "zero". Then the sum $\mathfrak{a} \oplus \mathfrak{n} \oplus \theta\mathfrak{n}$ is direct.

Now we prove that the sum $\mathfrak{k}+\mathfrak{a}+\mathfrak{n}$ is also direct. It is clear that $\mathfrak{a}\cap\mathfrak{n} = 0$ because $\mathfrak{a} \subseteq \mathfrak{f}_0$. Let $X \in \mathfrak{k}\cap(\mathfrak{a}\oplus\mathfrak{n})$. Then $\theta X = X$. But $\theta X \in \mathfrak{a} \oplus \theta\mathfrak{n}$. Thus $X \in \mathfrak{a} \oplus \mathfrak{n} \cap \mathfrak{a} \oplus \mathfrak{n}$ which implies $X \in \mathfrak{a}$. All this makes $X \in \mathfrak{p} \oplus \mathfrak{k}$ and $X = 0$.

Now we prove that $\mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n} = \mathfrak{f}$. An arbitrary $X \in \mathfrak{f}$ can be written as

$$X = H + X_0 + \sum_{\lambda\in\Sigma} X_\lambda$$

where $H \in \mathfrak{a}$, $X_0 \in \mathfrak{m}$ and $X_\lambda \in \mathfrak{f}_\lambda$. Now there are just some manipulations…

$$\sum_{\lambda\in\Sigma} X_\lambda = \sum_{\lambda\in\Sigma^+} (X_{-\lambda} + X_\lambda) = \sum_{\lambda\in\Sigma^+} (X_{-\lambda} + \theta X_{-\lambda}) + \sum_{\lambda\in\Sigma^+} (X_\lambda + \theta X_{-\lambda}), \tag{3.384}$$

but $\theta(X_{-\lambda} + \theta X_{-\lambda}) = X_{-\lambda} + \theta X_{-\lambda}$, then $X_{-\lambda} + X_{-\lambda} \in \mathfrak{k}$. Moreover, $X_\lambda, \theta X_{-\lambda} \in \mathfrak{f}_\lambda$, then $X_\lambda - \theta X_{-\lambda} \in \mathfrak{f}_\lambda \subseteq \mathfrak{n}$. Then

$$X = X_0 + \sum_{\lambda\in\Sigma^+} (X_{-\lambda} + \theta X_{-\lambda}) + H + \sum_{\lambda\in\Sigma^+} (X_\lambda - \theta X_{-\lambda}) \tag{3.385}$$

where the two first term belong to $\mathfrak{k}$, $H \in \mathfrak{a}$ and the last term belongs to $\mathfrak{n}$. □

Lemma 3.174.
There exists a basis $\{X_i\}$ of $\mathfrak{f}$ in which

(i) The matrices of $\operatorname{ad}\mathfrak{k}$ are symmetric,

(ii) The matrices of $\operatorname{ad}\mathfrak{a}$ are diagonal and real,

(iii) The matrices of $\operatorname{ad}\mathfrak{n}$ are upper triangular with zeros on the diagonal.

Proof. We have the orthogonal decomposition $\mathfrak{f} = \mathfrak{f}_0 \bigoplus_{\lambda\in\Sigma} \mathfrak{f}_\lambda$ given by proposition 3.172. Let $\{X_i\}$ be an orthogonal basis of $\mathfrak{f}$ compatible with this decomposition and in such an order that $i < j$ implies $\lambda_i \geqslant \lambda_j$. From the orthogonality of the basis it follows that the matrix of B_θ is diagonal. Thus the adjoint is the transposition.

(i) If $X \in \mathfrak{k}$, $(\operatorname{ad} X)^t = (\operatorname{ad} X)^* = -\operatorname{ad}\theta X = -\operatorname{ad} X$.

(ii) Each X_i is a restricted root; then $(\operatorname{ad} H)X_i = \lambda_i(H)X_i$, then the diagonal of $\operatorname{ad} H$ is made of $\lambda_i(H)$ whose are real.

(iii) If $Y_i \in \mathfrak{f}_{\lambda_i}$ with $\lambda_i \in \Sigma^+$, $(\operatorname{ad} Y_i)X_j$ has only components in $\mathfrak{f}_{\lambda_i+\lambda_j}$ with $\lambda_i+\lambda_j > \lambda_j$ because $\lambda_i \in \Sigma^+$. □

Lemma 3.175.
Let $\mathfrak{h}$ be a subalgebra of the real semisimple Lie algebra $\mathfrak{f}$. Then $\mathfrak{h}$ is a Cartan subalgebra if and only if $\mathfrak{h}^{\mathbb{C}}$ is Cartan in $\mathfrak{f}^{\mathbb{C}}$.

Proof. Direct sense. If $\mathfrak{h}$ is nilpotent in $\mathfrak{f}$, it is cleat that $\mathfrak{h}^{\mathbb{C}}$ is nilpotent in $\mathfrak{f}^{\mathbb{C}}$. We have to prove that $[x, \mathfrak{h}^{\mathbb{C}}] \subseteq \mathfrak{h}^{\mathbb{C}}$ implies $x \in \mathfrak{h}^{\mathbb{C}}$. As set, $\mathfrak{f}^{\mathbb{C}} = \mathcal{F} \oplus i\mathfrak{f}$ (but not as vector space !), then we can write $x = a + ib$ with a, $b \in \mathfrak{f}$. The assumption makes that for any $h \in \mathfrak{h}$, there exists $h', h'' \in \mathfrak{h}$ such that

$$[a + ib, h] = h + ih''.$$

This equation can be decomposed in $\mathfrak{f}$-part and $i\mathfrak{f}$-part: for any $h \in \mathfrak{h}$, there exists a $h' \in \mathfrak{h}$ such that $[a, h] = h'$, and for any $h \in \mathfrak{h}$, there exists a $h'' \in \mathfrak{h}$ such that $[b, h] = h''$. Thus a, $b \in \mathfrak{h}$ because $\mathfrak{h}$ is Cartan in $\mathfrak{f}$.
Inverse sense. The assumption is that $[x, \mathfrak{h}^{\mathbb{C}}] \subset \mathfrak{h}^{\mathbb{C}}$ implies $x \in \mathfrak{h}^{\mathbb{C}}$. In particular consider a $x \in \mathfrak{h}$ such that $[x, \mathfrak{h}] \subset \mathfrak{h}$. Then $x \in \mathfrak{h}^{\mathbb{C}}$ because $[x, \mathfrak{h}^{\mathbb{C}}] \subset \mathfrak{h}^{\mathbb{C}}$. But $\mathfrak{h}^{\mathbb{C}} \cap \mathfrak{f} = \mathfrak{h}$. □

In the complex case, the Cartan subalgebras all have same dimensions because they are maximal abelian.

3.12 Iwasawa decomposition of Lie groups

In this section, we show the main steps of the Iwasawa decomposition for a semisimple Lie group. For proofs, the reader will see [13] VI.4 and [3] III,§ 3,4 and VI,§ 3. In the whole section, G denotes a semisimple group, and $\mathfrak{g}$ its real (finite dimensional) Lie algebra. The two main examples that are widely used are $\mathrm{SL}(2,\mathbb{R})$ and $\mathrm{SO}(2,n)$.

3.12.1 Cartan decomposition

If $\mathfrak{g}$ is a finite dimensional Lie algebra and X, $Y \in \mathfrak{g}$, the composition of the adjoint $\operatorname{ad} X \circ \operatorname{ad} Y \colon \mathfrak{g} \to \mathfrak{g}$ makes sense.

Definition 3.176.
An involutive automorphism θ on a real *semi simple Lie algebra $\mathfrak{g}$ for which the form B_θ,*

$$B_\theta(X,Y) := -B(X,\theta Y) \tag{3.386}$$

(B is the Killing form on $\mathfrak{g}$) is positive definite is a ***Cartan involution****.*

Proposition 3.177.
There exists a Cartan involution for every real semisimple Lie algebra.

Problem and misunderstanding 16.
The theorem 4.1 in [3] is maybe a proof of this proposition.

See [3], theorem 4.1. Since $\theta^2 = id$, the eigenvalues of a Cartan involution are ± 1, and we can define the **Cartan decomposition** $\mathfrak{g}$

$$\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p} \tag{3.387}$$

into ± 1-eigenspaces of θ in such a way that $\theta = (-\operatorname{id})|_{\mathfrak{p}} \oplus \operatorname{id}|_{\mathfrak{k}}$. These eigenspaces are subject to the following commutation relations:

$$[\mathfrak{k},\mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{k},\mathfrak{p}] \subseteq \mathfrak{p}, \quad [\mathfrak{p},\mathfrak{p}] \subseteq \mathfrak{k}. \tag{3.388}$$

The dimension of a maximal abelian subalgebra of $\mathfrak{p}$ is the **rank** of $\mathfrak{g}$. One can prove that it does not depend on the choices (Cartan involution and maximal abelian subalgebra). We denote by $\mathfrak{a}$ such a maximal abelian subalgebras of $\mathfrak{p}$.

Lemma 3.178.
If $\mathfrak{g}_0$ is a real semisimple Lie algebra and θ a Cartan involution, then for all $X \in \mathfrak{g}_0$,

$$(\operatorname{ad} X)^* = -\operatorname{ad}(\theta X), \tag{3.389}$$

where the star on an operator on $\mathfrak{g}$ is defined by

$$B_\theta(X, AY) = B_\theta(A^* X, Y). \tag{3.390}$$

Lemma 3.179.
The set of operators $\operatorname{ad}(\mathfrak{a})$ is an abelian algebra whose elements are self-adjoint.

Proof. We have to prove that $(\operatorname{ad} H)^* = (\operatorname{ad} H)$ and $[\operatorname{ad} H, \operatorname{ad} I] = 0$ for every H, $I \in \mathfrak{a}$. First, note that $H \in \mathfrak{a} \subset \mathfrak{p}$, thus $\theta H = -H$, and $(\operatorname{ad} H)^* = -\operatorname{ad}(\theta H) = \operatorname{ad} H$.

For the second, $\operatorname{ad} H \circ \operatorname{ad} I = \operatorname{ad}(H \circ I)$ so that $[\operatorname{ad} H, \operatorname{ad} I] = \operatorname{ad}[H, I] = 0$ because $\mathfrak{a}$ is abelian. ☐

3.12.2 Root space decomposition

From the lemma, the operators $\operatorname{ad}(H)$ with $H \in \mathfrak{a}$ are simultaneously diagonalisable. That means that there exists a basis $\{X_i\}$ of $\mathfrak{g}$ and linear maps $\lambda_i \colon \mathfrak{a} \to \mathbb{R}$ such that

$$\operatorname{ad}(H) X_i = \lambda_i(H) X_i.$$

For each $\lambda \in \mathfrak{a}^*$, we define

$$\mathfrak{g}_\lambda = \{X \in \mathfrak{g} | (\operatorname{ad} H) X = \lambda(H) X, \forall H \in \mathfrak{a}\}. \tag{3.391}$$

Elements $0 \neq \lambda \in \mathfrak{a}^*$ such that $\mathfrak{g}_\lambda \neq 0$ are called **restricted roots** of $\mathfrak{g}$. The set of restricted roots is denoted by Σ.

Proposition 3.180.
The restricted root together with $\mathfrak{a}$ itself span the whole space:

$$\mathfrak{g} = \mathfrak{g}_0 \oplus_{\lambda\in\Sigma} \mathfrak{g}_\lambda, \tag{3.392}$$

This decomposition is called the ***restricted root space decomposition****.*

Proof. We first prove that the sum is direct. If the sum is not so, we can find a $H^* \in \mathfrak{g}_0$ and $X_i \in \mathfrak{g}_{\lambda_i}$ $(\lambda_i \in \Sigma)$ such that

$$H^* + \sum_i X_i = 0 \tag{3.393}$$

Let us consider

$$N = \{H \in \mathfrak{g}_0 |\text{ the } \lambda_i(H) \text{ are all differents and not zero}\}$$

A H which is not in N fulfils some relations as $\lambda_i(H) = \lambda_j(H)$ which are linear equations, so the complement of N is an union of hyperplanes and thus N is not empty. This allows us to consider a $H \in N$.

We have choice the X_i in $\mathfrak{g}_{\lambda_i}$, *i.e.*

$$(\operatorname{ad} A)X_i = \lambda_i(A)X_i \tag{3.394}$$

for all $A \in \mathfrak{a}$. In other words, X_i diagonalise $\operatorname{ad} A$ with eigenvalues $\lambda_i(A)$. Now, let us consider $\operatorname{ad} H$ for a $H \in N$. Since all the $\lambda_i(H)$ are different and not zero, the equation (3.394) implies that all the X_i (and H^*) are in separate eigenspaces of $\operatorname{ad} H$. Thus they are linearly independent, hence the equation (3.393) is not possible. The sum (3.392) is thus a direct sum. For the rest of the proof, see [3] theorem 4.2. □

Other properties of the root spaces are listed in the following proposition.

Proposition 3.181.
The spaces $\mathfrak{g}_{\lambda_i}$ satisfy also:

(i) $[\mathfrak{g}_\lambda, \mathfrak{g}_\mu] \subseteq \mathfrak{g}_{\lambda+\mu}$,

(ii) $\theta\mathfrak{g}_\lambda = \mathfrak{g}_{-\lambda}$*; in particular, if λ belongs to Σ, then $-\lambda$ belongs to Σ too,*

(iii) $\mathfrak{g}_0 = \mathfrak{a} \oplus \mathcal{Z}_{\mathfrak{k}}(\mathfrak{a})$ *orthogonally.*

Problem and misunderstanding 17.
Il faut définir quelque part ce qu'est cet espace $\mathcal{Z}_{\mathfrak{k}}(\mathfrak{a})$

3.12.3 Positivity, convex cone and partial ordering

Definition 3.182.
Let V be a vector space. A ***positivity notion*** *(see [13], page 154) is the data of a subset V^+ of V such that*

(i) for every nonzero $v \in V$, we have $v \in V^+$ xor $-v \in V^+$,

(ii) for every v, $w \in V^+$ and every $\mu \in \mathbb{R}^+$, the elements $v + w$ and μv are positive.

If $v \in V^+$, we say that v is **positive** and we note $v > 0$.

Definition 3.183.
A ***convex cone*** *in a vector space A is a subset C such that*

(i) $x \in C$ and $t \in \mathbb{R}^+$ imply $tx \in C$,

(ii) $x, y \in C$ implies $x + y \in C$,

(iii) $C \cap (-C) = \{0\}$.

To a convex cone C is attached a notion of positivity by defining $x \geqslant 0$ if and only if $x \in C$. The converse is also true: if we have a notion of positivity on V, we define the corresponding convex cone by

$$V^+ = \{x \in V \text{ st } x \geqslant 0\}. \tag{3.395}$$

A **linear partial ordering relation** is a partial ordering $\leqslant$ such that

- $A \leqslant B$ implies $A + C \leqslant B + C$ for all C,
- $\lambda A \leqslant \lambda B$ for all $\lambda \in \mathbb{R}^+$.

From a positivity notion gives rise to a linear partial ordering on V by defining $x \geqslant y$ if and only if $y - x \geqslant 0$.

3.12.4 Iwasawa decomposition

Let us consider a notion of positivity on $\mathfrak{a}^*$ and denote by Σ^+ the set of positive roots. We define

$$\mathfrak{n} := \oplus_{\lambda\in\Sigma^+}\mathfrak{g}_\lambda. \tag{3.396}$$

The **Iwasawa decomposition** is given by the following theorem ([13], theorem 5.12):

Theorem 3.184.
Let G be a linear connected semisimple group and $A = \exp\mathfrak{a}$, $N = \exp\mathfrak{n}$ where $\mathfrak{a}$ and $\mathfrak{n}$ are the previously defined algebras. Then A, N and AN are simply connected subgroups of G and the multiplication map

$$\begin{aligned} \phi\colon A\times N\times K &\to G\\ (a,n,k) &\mapsto ank \end{aligned} \tag{3.397}$$

is a global analytic diffeomorphism. In particular, the Lie algebra $\mathfrak{g}$ decomposes as vector space direct sum

$$\mathfrak{g} = \mathfrak{a}\oplus\mathfrak{n}\oplus\mathfrak{k}. \tag{3.398}$$

*The group AN is a solvable subgroup of G which is called the **Iwasawa group**, or Iwasawa component of G.*

Remark 3.185.
It can be proved that this theorem is independent of the choices: the Cartan involution, the maximal abelian subalgebra $\mathfrak{a}$ and the notion of positivity.

Notice that A, N and K are unique up to isomorphism. Their matricial representation of course depend on choices.

This theorem from [3], chapter VI, Theorem 3.4. will be useful.

Theorem 3.186.
The Lie algebra $\mathfrak{a}\oplus\mathfrak{k}$ is solvable.

This theorem implies that the group AN is solvable.[25] Before to go into concrete situations, let us prove an useful property of the $\mathfrak{k}$ part of $\mathfrak{g}$:

Theorem 3.187.

$$\mathcal{S}tab(\mathfrak{k}) = K$$

for the adjoint action of G on $\mathfrak{k}$.

The proof of it is given by two lemmas. [25]

Lemma 3.188.
For any $k\in K$,

$$\mathrm{Ad}(k)\mathfrak{k} = \mathfrak{k},$$

and

Lemma 3.189.
If for any $L\in\mathfrak{k}$, $\mathrm{Ad}(x)L$ belongs to $\mathfrak{k}$, then $x\in K$.

Proof of lemma 3.188. Let us take a $L\in\mathfrak{k}$ and define $M\in K$ $k = e^M$. We have $\mathrm{Ad}(k)L = e^{\mathrm{ad}\,M}L$. But in general, we have the relations (3.388) which give $e^{\mathrm{ad}\,M}L\in\mathfrak{k}$. Then $\mathrm{Ad}(k)\mathfrak{k}\subset\mathfrak{k}$.

In order to show that $\mathfrak{k}\subset\mathrm{Ad}(k)\mathfrak{k}$, let us consider a $L\in\mathfrak{k}$. We have to find a $N\in\mathfrak{k}$ such that $\mathrm{Ad}(k)N = L$. It is clear that $N = \mathrm{Ad}(k^{-1})L$ fulfils the conditions. □

Proof of lemma 3.189. Let us consider $X\in\mathfrak{g}$ such that $x = e^X$. We have $e^{\mathrm{ad}\,X}L\in\mathfrak{k}$ for all $L\in\mathfrak{k}$. This implies that all the terms of the expansion of $e^{\mathrm{ad}\,X}L$ are in $\mathfrak{k}$. In particular, $[X,L]\in\mathfrak{k}$ for all $L\in\mathfrak{k}$. Let us consider the Cartan decomposition of X : $X = X_k + X_p$. We need X such that

$$[X_k, L] + [X_p, L] \in\mathfrak{k}$$

for any $L\in\mathfrak{k}$. But inclusions (3.388) make $[X_p, L]\in\mathfrak{p}$. Then the X_p part of X must vanish (because $\mathfrak{g} = \mathfrak{k}\oplus\mathfrak{p}$ is a direct sum). □

[25] J'esère que ce que je raconte ici n'est pas trop débile pcq j'ai pas été fouiller à fond.

3.13 Representations

Source :[14]

We are interested in the adjoint representation on a common vector space; we will not discuss the importance of some more complicated features as the "locally convex" condition. We only mention it.

Definition 3.190.
*If V is a locally convex space, a **continuous representation** of a Lie group G on V is a left invariant action $\pi\colon G\times V\to V$ such that for any $x\in G$, the map $\pi(x)\colon V\to V$ is a linear endomorphism of V.*

If $\mathfrak{g}$ is a Lie algebra, a **representation** of $\mathfrak{g}$ in V is a bilinear map $\sigma\colon V\to V$ such that

$$\sigma([X,Y])v=[\sigma(X),\sigma(Y)]v \tag{3.399}$$

where the second bracket is the usual commutator of operators. In other words, $\sigma\colon\mathfrak{g}\to\operatorname{End}V$ is an algebra homomorphism.

A vector space equipped with a representation of a Lie algebra $\mathfrak{g}$ is a **$\mathfrak{g}$-module**. A *complete* locally convex space equipped with a representation of a Lie group is a **G-module**.

Let us write down Schur's lemma:

Lemma 3.191.
If $\phi\colon\mathfrak{g}\to\mathfrak{gl}(V)$ is irreducible, then the only endomorphism of V which commutes with all $\phi(\mathfrak{g})$ are multiples of identity.

If π is a representation of G in a (eventually complex) vector space V, an **invariant subspace** is a vector subspace $W\subset V$ such that $\pi(x)W\subset W$ for any $x\in G$. A continuous representation in a complete locally compact vector space V is **irreducible** if $\{0\}$ and V are the only two invariant closed subspaces of V.

In the case of finite dimensional vector space, any subspace is closed; in this class, we find back the usual notion of irreducibility.

An **unitary** representation of G is a continuous representation π of G in a complex (or real) Hilbert space H such that $\pi(x)$ is unitary for any $x\in G$. This is: π is unitary if and only if $\forall x\in G$, v, $w\in H$,

$$\langle\pi(x)v,w\rangle=\langle v,\pi(x)^{-1}w\rangle. \tag{3.400}$$

A continuous and finite dimensional representation is **unitarisable** if there exists an hermitian product for which the representation is unitary.

Now a great proposition without proof:

Proposition 3.192.
Let G be a compact Lie group[26]*. Then every representation on a finite dimensional vector space is unitarisable.*

3.14 Other results about Cartan algebras

Lemma 3.193.
A Cartan subalgebra of a semisimple complex Lie algebra is maximally abelian.

Proof. If $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$, proposition 3.73 provides $H_0\in\mathfrak{g}$ such that $\mathfrak{h}=\mathfrak{g}_0(H_0)$; in particular $H_0\in\mathfrak{h}$. We are going to prove that if $H_1,H_2\in\mathfrak{h}$, then for every $Y\in\mathfrak{g}$ we have $B\big([H_1,H_2],Y\big)=0$, so that the non degeneracy of the Killing form will conclude that $[H_1,H_2]=0$.

Let $X\in\mathfrak{g}(H_0,\lambda)$, $H\in\mathfrak{h}$. The map $\operatorname{ad}X\circ\operatorname{ad}H$ sends $\mathfrak{g}(H_0,\mu)$ to $\mathfrak{g}(H_0,\lambda+\mu)$. If we choose a basis of $\mathfrak{g}$ made up with basis of the spaces $\mathfrak{g}(H_0,\lambda_i)$ (by the primary decomposition theorem) it is clear that $B(H,X)=\operatorname{Tr}(\operatorname{ad}H\circ\operatorname{ad}X)=0$. In particular with $H=[H_1,H_2]$ we get $B\big([H_1,H_2],X\big)=0$.

On the other hand, $\mathfrak{h}$ is solvable because it is nilpotent. Since the adjoint action provides a representation of $\mathfrak{h}$ on $\mathfrak{h}$, corollary 3.21 says that we have basis of $\mathfrak{h}$ in which all the matrices of are upper triangular. Now if A, B and C are upper triangular matrices, ABC and BAC have same elements on the diagonal;in particular they traces are the equal: $\operatorname{Tr}(ABC)=\operatorname{Tr}(BAC)$. Let us consider H_1, H_2, $H\in\mathfrak{h}$ By Jacobi, $\operatorname{ad}[H_1,H_2]=[\operatorname{ad}H_1,\operatorname{ad}H_2]$, then

$$\begin{aligned}\operatorname{Tr}(\operatorname{ad}[H_1,H_2]\operatorname{ad}H)&=\operatorname{Tr}(\operatorname{ad}H_1\operatorname{ad}H_2\operatorname{ad}H)-\operatorname{Tr}(\operatorname{ad}H_2\operatorname{ad}H_1\operatorname{ad}H)\\&=\operatorname{Tr}(\operatorname{ad}H_2\operatorname{ad}H_1\operatorname{ad}H)-\operatorname{Tr}(\operatorname{ad}H_1\operatorname{ad}H_2\operatorname{ad}H)\\&=0.\end{aligned} \tag{3.401}$$

[26] Verifie si il faut que ce soit de Lie

Up to now we had seen that $B([H_1,H_2],H)=0$ and $B(H,X)=0$ if $X\in\oplus_{\lambda\neq 0}\mathfrak{g}(H_0,\lambda)$. In the latter, we can consider $[H_1,H_2]$ as H. Then

$$B([H_1,H_2],Y)=0$$

for all $Y\in\mathfrak{g}$. Then $[H_1,H_2]=0$ because the Killing form is nondegenerate ($\mathfrak{g}$ is semisimple). This proves that $\mathfrak{h}$ is abelian.

Now it remains to see that $\mathfrak{h}$ is contained in no larger abelian subalgebra of $\mathfrak{g}$. For this, we naturally consider a larger abelian subalgebra $\mathfrak{h}'$ of $\mathfrak{g}$. For any $H'\in\mathfrak{h}'$ and $H\in\mathfrak{h}$, we have $[H,H']=0$. In particular $[H',H_0]=0$; the property

$$\mathfrak{h}=\mathfrak{g}(H_0,0)=\{X\in\mathfrak{g}\text{ st }(\operatorname{ad}H_0)^kX=0\text{ for a cerain }k\in\mathbb{N}\}.$$

makes $H'\in\mathfrak{h}$. □

Proposition 3.194.
Let $\mathfrak{g}$ be a Lie algebra, $x\in\mathfrak{g}$ and

$$\mathfrak{g}^x=\{y\in\mathfrak{g}\ st\ \exists n\in\mathbb{N}:(\operatorname{ad}x)^ny=0\}. \tag{3.402}$$

Then $\mathfrak{g}^x$ is a subalgebra of $\mathfrak{g}$ which is its own centralizer in $\mathfrak{g}$.

Proof. Since $\operatorname{ad}(x)$ is a derivation of $\mathfrak{g}$ (cf. 3.1),

$$(\operatorname{ad}x)^n([u,v])=\sum_{k=0}^n\binom{n}{k}[(\operatorname{ad}x)^ku,(\operatorname{ad}x)^{n-k}v];$$

then $[\mathfrak{g}^x,\mathfrak{g}^x]\subset\mathfrak{g}^x$. This proves that $\mathfrak{g}^x$ is a subalgebra of $\mathfrak{g}$. Let $y\in\mathfrak{g}$ be such that $[y,\mathfrak{g}^x]\subset\mathfrak{g}^x$. Clearly $[x,y]\in\mathfrak{g}^x$ (because $x\in\mathfrak{g}^x$) then $(\operatorname{ad}x)^ny=(\operatorname{ad}x)^{n-1}[x,y]$, so that $y\in\mathfrak{g}^x$. □

Lemma 3.195.
If $A\colon V\to V$ is a linear operator on a finite dimensional vector space, then there exists a positive integer p such that $A^p(V)=A^{p+1}(V)$.

Proof. We build a basis of V in the following manner. Since $A(V)$ is a subspace of V, we can begin our basis with $\{Y_i\}$, a basis of the component of $A(V)$ in V. Next, $A^2(V)$ is a subspace of $A(V)$, then we can consider $\{X_i^1\}$, a basis of the vector space $A(V)\backslash A^2(V)$, and so on... $\{X_i^p\}$ are vectors in $A^n(V)$ but not in $A^{n+1}V$. Since the vector space has only a finite number of basis vectors, there is a p such that $\{X_i^p\}=\emptyset$. □

Now we consider $W=\{u\in V\text{ st }\exists n\in\mathbb{N}:A^nu=0\}$ and $v\in V$. There exists a $v'\in V$ such that $A^p(v)=A^{p+1}(v')$. Writing $v=A(v')+(v-A(v'))$, we find

$$V\subset A(V)+W \tag{3.403}$$

because $A^p(v)-A^{p+1}(v')=0$, $v-A(v')\in W$.

If we apply A on this, we find $A(V)\subset A^2(V)+A(W)$. Reinserting it into the right hand side of (3.403), we find $V\subset A^2(V+W)$ and repeating p times this process, we find $V=A^p(V)+W$ and the sum is direct because none of the elements of $A^p(V)$ is annihilated by A:

$$V=A^p(V)\oplus W. \tag{3.404}$$

Proposition 3.196.
Let $\mathfrak{g}$ be a Lie algebra and $x\in\mathfrak{g}$. Then there exists a subspace $\mathfrak{g}_x$ of $\mathfrak{g}$ such that $\mathfrak{g}=\mathfrak{g}_x\oplus\mathfrak{g}^x$ and $[\mathfrak{g}^x,\mathfrak{g}_x]\subset\mathfrak{g}_x$.

Proof. We claim that the space is given by

$$\mathfrak{g}_x=(\operatorname{ad}x)^p\mathfrak{g} \tag{3.405}$$

where p is taken large enough to have $(\operatorname{ad}x)^p\mathfrak{g}=(\operatorname{ad}x)^{p+1}\mathfrak{g}$. The lemma and the discussion below show the correctness of the definition of $\mathfrak{g}_x$ and that $\mathfrak{g}=\mathfrak{g}_x\oplus\mathfrak{g}^x$. It remains to be proved that $[\mathfrak{g}^x,\mathfrak{g}_x]\subset\mathfrak{g}_x$. For we will prove (by induction with respect to m) for any m that $(\operatorname{ad}x)^my=0$ implies $(\operatorname{ad}y)\mathfrak{g}_x\subset\mathfrak{g}_x$.

For $m=1$, the induction assumption becomes $[x,y]=0$ and Jacobi gives $\operatorname{ad}x\circ\operatorname{ad}y=\operatorname{ad}y\circ\operatorname{ad}x$, then $(\operatorname{ad}y)\mathfrak{g}_x=(\operatorname{ad}x)^p(\operatorname{ad}y)\mathfrak{g}\subset\mathfrak{g}_x$. Now we suppose that $(\operatorname{ad}x)^{m-1}z=0$ implies $(\operatorname{ad}z)\mathfrak{g}_x\subset\mathfrak{g}_x$ and we consider $y\in\mathfrak{g}$ such that $(\operatorname{ad}x)^my=0$ and $u\in\mathfrak{g}_x$. We are going to show that $(\operatorname{ad}y)u\in\mathfrak{g}_x$. Let f be the characteristic polynomial of $\operatorname{ad}x$:

$$f(t)=\det\left(\operatorname{ad}x-t\mathbb{1}\right)$$

where $\operatorname{ad}x$ and $\mathbb{1}$ are taken on $\mathfrak{g}_x$. Since $(\operatorname{ad}x)u=0$, $f(0)\neq 0$ and by the Cayley-Hamilton theorem, $f(\operatorname{ad}x)u=0$. Then

$$(f(\operatorname{ad}x)\operatorname{ad}y)u=(f(\operatorname{ad}x)\operatorname{ad}y-(\operatorname{ad}y)f(\operatorname{ad}x))u, \tag{3.406}$$

and, on the other hand, $\forall q\in\mathbb{N}$,

$$(\operatorname{ad}x)^q\operatorname{ad}y-\operatorname{ad}y(\operatorname{ad}x)^q=\sum_{r=0}^{q-1}(\operatorname{ad}x)^r(\operatorname{ad}[x,y])(\operatorname{ad}x)^{q-r-1}.$$

It follows that $f(\operatorname{ad}x)\operatorname{ad}y-(\operatorname{ad}y)f(\operatorname{ad}x)$ is a linear combination of terms of the form

$$(\operatorname{ad}x)^a(\operatorname{ad}[x,y])(\operatorname{ad}x)^b$$

and the induction hypothesis shows that $f(\operatorname{ad}x)(\operatorname{ad}y)u\in\mathfrak{g}_x$.

Now we consider a n such that $(\operatorname{ad}x)^n\mathfrak{g}^x=0$; the fact that $f(0)\neq 0$ implies the existence of polynomials $g(t)$ and $h(t)$ such that $g(t)t^n+h(t)f(t)=1$. If we decompose $(\operatorname{ad}y)u=v+w$ with respect to $\mathfrak{g}=\mathfrak{g}_x\oplus\mathfrak{g}^x$ we find

$$\begin{aligned}(\operatorname{ad}y)u&=[g(\operatorname{ad}x)(\operatorname{ad}x)^n+h(\operatorname{ad}x)f(\operatorname{ad}x)](\operatorname{ad}y)u\\&=f(\operatorname{ad}x)(\operatorname{ad}x)^nv+h(\operatorname{ad}x)f(\operatorname{ad}x)(\operatorname{ad}y)u\in\mathfrak{g}_x.\end{aligned}\tag{3.407}$$

□

Proposition 3.197.
Let $\mathfrak{g}$ be a Lie algebra and $x\in\mathfrak{g}$ such that $\mathfrak{g}^x$ is as small as possible. Then $\mathfrak{g}^x$ is a Cartan subalgebra.

Proof. From proposition 3.194, it is sufficient to prove that $\mathfrak{g}^x$ is nilpotent. Let $y\in\mathfrak{g}^x$ and $f_y(t)$ be the characteristic polynomial of $\operatorname{ad}y$. Since it is a subalgebra, $\mathfrak{g}^x$ is stable under $\operatorname{ad}y$ and proposition 3.196 makes $\mathfrak{g}_x$ also stable under $\operatorname{ad}y$. Then $\operatorname{ad}y$ can be written under a bloc-diagonal form with respect to the decomposition $\mathfrak{g}=\mathfrak{g}_x\oplus\mathfrak{g}^x$, so that the characteristic polynomial can be factorised as

$$f_y(t)=g_y(t)h_y(t)\tag{3.408}$$

where g_y and h_y are the characteristic polynomials of the restrictions of $\operatorname{ad}y$ to $\mathfrak{g}^x$ and $\mathfrak{g}_x$. Let $(y_1,\ldots,y_m)$ be a basis of $\mathfrak{g}^x$ and t^n, the greatest power of t which divide all the $g_y(t)$ with $y\in\mathfrak{g}^x$. The coefficient of t^n in $g_{c^iy_i}(t)$ is a polynomial with respect to the c^i because of the expression

$$g_{c^iy_i}(t)=\det\Big(\operatorname{ad}(c^iy_i)-t\mathbb{1}\Big).$$

Let u be this polynomial: $g_{c^iy_i}(t)=\ldots+u(c^1,\ldots,c^m)t^n$. By definition of n, this is not an identically zero polynomial and there are no terms with t^{n-1}. For the same reasons, we have a polynomial v such that

$$h_{c^iy_i}(0)=v(c^1,\ldots,c^m).\tag{3.409}$$

We know that none of the non-zero elements in $\mathfrak{g}_x$ are annihilated by $\operatorname{ad}x$ (because of the definition of $\mathfrak{g}^x$). Then $h_x(0)\neq 0$ and v is not identically zero. With all this we can find some $c^i\in\mathbb{C}$ such that $u(c^1,\ldots,c^m)v(c^1,\ldots,c^m)\neq 0$. If $y=c^iy_i$, the coefficient of t^n in $f_y(t)$ is $u(c)v(c)\neq 0$, so that $f_y(t)$ is not divisible by t^{n+1}.

But in the other hand $\mathfrak{g}^x$ has minimal dimension, then $\dim\mathfrak{g}^y\geqslant m=\dim\mathfrak{g}^x$. Moreover $t^{\dim\mathfrak{g}^y}$ divide $f_y(t)$ because there is a certain power of $\operatorname{ad}y$ which has zero as eigenvalue with multiplicity $\dim\mathfrak{g}^y$[27]. Since $f_y(t)$ can not be divided by t^{n+1} this shows that $n+1>\dim\mathfrak{g}^y$ and $n\geqslant\dim\mathfrak{g}^y\geqslant m$.

Now we consider y, any element of $\mathfrak{g}^x$. From the fact that t^n divide all the $g_y(t)$ and that $n\geqslant m$, we see that t^m divide $g_y(t)$. But the degree of $g_y(t)$ is $\dim\mathfrak{g}^x=m$. Finally, $g_y(t)=m$ and $\operatorname{ad}y$ is nilpotent on $\mathfrak{g}^x$ for any $y\in\mathfrak{g}^x$.

The Engel's theorem 3.32 makes $\mathfrak{g}^x$ nilpotent. □

The following holds for complex or real Lie algebras and comes from [15] see also [8]. We denote by $\mathbb{K}$ the base field of $\mathfrak{g}$, i.e. $\mathbb{R}$ or $\mathbb{C}$. For $X\in\mathfrak{g}$ and $\lambda\in\mathbb{K}$ we define

$$\mathfrak{g}(X,\lambda)=\{Y\in\mathfrak{g}\text{ st }(\operatorname{ad}X-\lambda\mathbb{1})^nY=0\text{ for a certain }n\in\mathbb{N}\}.\tag{3.410}$$

A first useful result is given in

Lemma 3.198.
If $Z\in\mathfrak{g}$, then

$$[\mathfrak{g}(Z,\lambda),\mathfrak{g}(Z,\mu)]\subset\mathfrak{g}(Z,\lambda+\mu),$$

in particular $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$.

[27] This is not a good reason.

Proof. We consider $X_\lambda \in \mathfrak{g}(Z,\lambda)$ and $X_\mu \in \mathfrak{g}(Z,\mu)$. We have

$$\begin{aligned}\big(\operatorname{ad} Z-(\lambda+\mu)I\big)[X_\lambda,X_\mu] = [&(\operatorname{ad} Z-\lambda I)X_\lambda,X_\mu]\\ &+[X_\lambda,(\operatorname{ad} Z-\mu I)X_\mu].\end{aligned} \tag{3.411}$$

By induction,

$$\big(\operatorname{ad} Z-(\lambda+\mu)\mathbb{1}\big)^n[X_\lambda,X_\mu] = \sum_{i=0}^{\infty}\binom{n}{i}[(\operatorname{ad} Z-\lambda I)^iX_\lambda,(\operatorname{ad} Z-\mu I)^{n-i}X\mu]. \tag{3.412}$$

It will become zero for large enough n. □

An element $X\in\mathfrak{g}$ is **regular** if $\dim\mathfrak{g}(X,0)$ is minimum[28]. This minimum is the **rank** of $\mathfrak{g}$.

Proposition 3.199.
If $X\in\mathfrak{g}$ is a regular element then the algebra

$$\mathfrak{h}=\mathfrak{g}(X,0)=\{Y\in\mathfrak{g}\textit{ st }(\operatorname{ad} X)^nY=0\textit{ for some }n\in\mathbb{N}\} \tag{3.413}$$

is nilpotent.

Proof. We have to show that for any $H\in\mathfrak{h}$, the endomorphism $\operatorname{ad} H$ of $\mathfrak{h}$ is nilpotent. Consider the characteristic polynomial of $\operatorname{ad} X$

$$p(t)=\det(\operatorname{ad} X-t\mathbb{1})=t^rq(t)$$

where t^r is the maximal factorization of t; in other words, $q(t)$ is not divisible by t and $r=\dim\mathfrak{h}$. In particular

$$\mathfrak{h}=\{Y\in\mathfrak{g}\text{ st }(\operatorname{ad} X)^rY=0\}. \tag{3.414}$$

Let

$$\mathfrak{k}=\{Y\in\mathfrak{g}\text{ st }q(\operatorname{ad} X)Y=0\} \tag{3.415}$$

From the Cayley-Hamilton theorem (A.2), $p(\operatorname{ad} X)=0$, then $(\operatorname{ad} X)^rq(\operatorname{ad} X)=0$ and $\mathfrak{g}=\mathfrak{h}\oplus\mathfrak{k}$. Moreover $\mathfrak{h}$ and $\mathfrak{k}$ are $\operatorname{ad} X$-invariants: $(\operatorname{ad} X)\mathfrak{h}\subseteq\mathfrak{h}$ and $(\operatorname{ad} X)\mathfrak{k}\subseteq\mathfrak{k}$.

Every weight of $\operatorname{ad} X$ are in $\mathbb{C}$. As we know that $\mathfrak{h}$ is Cartan in $\mathfrak{g}$ if and only if $\mathfrak{h}^{\mathbb{C}}$ is Cartan in $\mathfrak{g}^{\mathbb{C}}$, we can suppose that $\mathfrak{g}$ is a complex algebra by considering $\mathfrak{g}^{\mathbb{C}}$ if $\mathfrak{g}$ is real. So all the weight are in the base field and we can define

$$\mathfrak{k}=\sum_{\lambda\in\Delta}\mathfrak{g}(X,\lambda).$$

where Δ is the set of all the non zero weight of $\operatorname{ad} X$. A property[29] of the weight space is that

$$\mathfrak{g}=\mathfrak{g}(X,\lambda_1)\oplus\ldots\oplus\mathfrak{g}(X,\lambda_m)$$

if the λ_i's are the weight of $\operatorname{ad} X$. Now we prove that $\sum_{\lambda\neq0}\mathfrak{g}(X,\lambda)=\mathfrak{k}$. First consider a $Y\in\mathfrak{g}(X,\lambda)$ which can be decomposed as $Y=H+K$ with $H\in\mathfrak{h}$ and $K\in\mathfrak{k}$. Then $(\operatorname{ad} X-\lambda\mathbb{1})^n(H+K)=(\operatorname{ad} X-\lambda\mathbb{1})^nH+(\operatorname{ad} X-\lambda\mathbb{1})^nK$ where the first term is not zero (because $H\in\mathfrak{h}$) and lies in $\mathfrak{h}$ while the second term lies in $\mathfrak{k}$. Then the sum can be zero only if $H=0$.

□

Let $\mathfrak{g}$ be a complex semisimple Lie algebra, $H\in\mathfrak{g}$ and $0=\lambda_0,\lambda_1,\ldots,\lambda_r$, the eigenvalues of $\operatorname{ad} H$. For any $\lambda\in\mathbb{C}$, one can consider

$$\mathfrak{g}(H,\lambda)=\{X\in\mathfrak{g}\text{ st }(\operatorname{ad} H-\lambda I)^kX=0\}. \tag{3.416}$$

From the Jordan decomposition, $g(H,\lambda)=0$ except if λ is one of the λ_i, and

$$\mathfrak{g}=\bigoplus_{i=0}^{r}g(H,\lambda_i). \tag{3.417}$$

An element $H\in\mathfrak{g}$ is **regular** if

$$\dim g(H,0)=\min_{X\in\mathfrak{g}}\dim g(X,0).$$

Let H_0 be a regular element and $\mathfrak{h}=\mathfrak{g}(H_0,0)$.

[28]Anglais ?
[29]Que je dois encore faire, cf Sagle

Lemma 3.200.
The algebra $\mathfrak{h} = \mathfrak{g}(H_0, 0)$ *is nilpotent*

Proof. Let $0 = \lambda_0, \lambda_1, \ldots, \lambda_r$ be the eigenvalues of $\operatorname{ad} H_0$ and

$$\mathfrak{g}' = \sum_{i=1}^{r} \mathfrak{g}(H_0, \lambda_i)$$

which is a subspace of $\mathfrak{g}$. From the lemma,

$$[\mathfrak{g}(H_0, 0), \mathfrak{g}(H_0, \lambda_i)] \subset \mathfrak{g}(H_0, \lambda_i) \subset \mathfrak{g}'.$$

For each $H \in \mathfrak{h}$, we denote H', the restriction of $\operatorname{ad} H$ to $\mathfrak{g}'$ and $d(H) = \det H'$. The function $H \to d(H)$ is a polynomial on $\mathfrak{h}$ in the sense of the coordinates on $\mathfrak{h}$ as vector space. If H_0' has a zero eigenvalue we would have $\operatorname{ad}(H_0)X = 0$ for some $X \in \mathfrak{g}'$. In this case $[H_0, X] = 0$, but $X \in \mathfrak{g}(H_0, \lambda_i)$, then for a certain k, $(\operatorname{ad} H_0 - \lambda_i)^k X = 0$, so that $\lambda_i X = 0$. Since $\mathfrak{g}$ is defined by excluding λ_0, $X = 0$. Thus H_0' has only non zero eigenvalues and $d(H_0) \neq 0$.

We know that a polynomial which is zero on an open set is identically zero; then on any open set of $\mathfrak{h}$, d has a non zero value somewhere. In particular,

$$S = \{H \in \mathfrak{h} \text{ st } d(H) \neq 0\}$$

is dense in $\mathfrak{h}$. We consider a $H \in S$. The endomorphism H' has only non zero eigenvalues, so that $\mathfrak{g}(H, 0) \subset \mathfrak{h}$ from lemma 3.198; but H_0 is regular, then $\mathfrak{g}(H, 0) \subset \mathfrak{h}$. Thus the restriction of $\operatorname{ad} H$ to $\mathfrak{h}$ is nilpotent because it is nilpotent on $\mathfrak{g}(H, 0)$[30].

If $l = \dim \mathfrak{h}$, then $(\operatorname{ad}_{\mathfrak{h}} H)^l = 0$ because $\operatorname{ad}_{\mathfrak{h}} H$ is nilpotent. By continuity, this equation is true for any $H \in \mathfrak{h}$ from the density of S in $\mathfrak{h}$. Then $\mathfrak{h}$ is nilpotent.

□

Here is an alternative proof (that I do not really understand) for theorem 3.77.

Theorem 3.201.
Let $\mathfrak{g}$ *be a complex Lie algebra with Cartan subalgebra* $\mathfrak{h}$*. Then* $\mathfrak{g}_0 = \mathfrak{h}$*.*

Proof. Since $\mathfrak{h}$ is Cartan, it is nilpotent. So $\mathfrak{h} \subset \mathfrak{g}_0$. If $v \in \mathfrak{g}_0$, there exists a n such that for any $z \in \mathfrak{h}$, $(\operatorname{ad} z)^n v = 0$. The fact that $\mathfrak{h}$ is nilpotent makes $(\operatorname{ad} z_n) \circ \ldots \circ (\operatorname{ad} z_1)v = 0$ for any $z \in \mathfrak{g}_0$ and for all $z_1, \ldots, z_n \in \mathfrak{h}$. If we write $(\operatorname{ad} z_1)v$ with $v \in \mathfrak{g}_0 \backslash \mathfrak{h}$, we can always choose z_1 in order the result to *not* be $\mathfrak{h}$. Next we can choose $z_2 \in \mathfrak{h}$ such that $(\operatorname{ad} z_2) \circ (\operatorname{ad} z_1)v$ is also not in $\mathfrak{h}$ and so on... Since $\mathfrak{g}_0$ is nilpotent, we always finish on zero. If n is the maximum of adjoint that we can take before to fall into zero; we have

$$[h, (\operatorname{ad} z_{n-1}) \circ (\operatorname{ad} z_1)v] = 0$$

for all $h \in \mathfrak{h}$ and with a good choice of z_i, it contradicts the fact that $\mathfrak{h}$ is Cartan. □

3.15 Universal enveloping algebra

Let $\mathcal{A}$ be a Lie algebra. One knows that the composition law $(X, Y) \to [X, Y]$ is often non associative. In order to build an associative Lie algebra which "looks like" $\mathcal{A}$, one considers $T(\mathcal{A})$, the tensor algebra of $\mathcal{A}$ (as vector space) and $\mathcal{J}$ the two-sided ideal in $T(\mathcal{A})$ generated by elements of the form

$$X \otimes Y - Y \otimes X - [X, Y]$$

for $X, Y \in \mathcal{A}$. The **universal enveloping algebra** of $\mathcal{A}$ is the quotient

$$U(\mathcal{A}) = T(\mathcal{A})/\mathcal{J}. \tag{3.418}$$

For $X \in \mathcal{A}$, we denote by X^* the image of X by canonical projection $\pi\colon T(\mathcal{A}) \to U(\mathcal{A})$ and by 1 the unit in $U(\mathcal{A})$. One has $1 \neq 0$ if and only if $\mathcal{A} \neq \{0\}$.

A property without proof[31] (see [3] page 90):

[30] Ce paragraphe n'est pas vraiment clair...
[31] La preuve est à partir de 21# de Lie

Proposition 3.202.
Le V be a vector space on K. Then there is a natural bijection between the representations of $\mathcal{A}$ on V and the ones of $U(\mathcal{A})$ on V. If ρ is a representation of $\mathcal{A}$ on V, the corresponding ρ^* of $U(\mathcal{A})$ is given by

$$\rho(X) = \rho^*(X^*)$$

($X \in \mathcal{A}$).

Let $\{X_1, \ldots, X_n\}$ be a basis of $\mathcal{A}$. For a n-uple of complex numbers (t_i), one defines

$$X^*(t) = \sum_{i=1}^{n} t_i X_i^*. \tag{3.419}$$

On the other hand, we consider a n-uple of positive integers $M = (m_1 + \ldots m_n)$, and the notation

$$\begin{aligned} |M| &= m_1 + \ldots + m_n \\ t^M &= t_1^{m_1} \cdots t_n^{m_n}. \end{aligned} \tag{3.420}$$

When $|M| > 0$, we denote by $X^*(M) \in U(\mathcal{A})$ the coefficient of t^M in the expansion of $(|M|!)^{-1}(X^*(t))^{|M|}$. If $|M| = 0$, the definition is $X^*(0) = 1$. Once again a proposition without proof[32]:

Proposition 3.203.
The smallest vector subspace of $U(\mathcal{A})$ which contains all the elements of the form $X^*(M)$ is $U(\mathcal{A})$ itself:

$$U(\mathcal{A}) = \mathrm{Span}\{X^*(M) : M \in \mathbb{N}^n\}.$$

Corollary 3.204.
Let $\mathcal{A}$ be a Banach algebra of dimension n, $\mathcal{B}$ a Banach subalgebra of dimension $n-r$ and a basis $\{X_1, \ldots, X_n\}$ of $\mathcal{A}$ such that the $n-r$ last basis vectors are in $\mathcal{B}$. We denotes by B the vector subspace of $U(\mathcal{A})$ spanned by the elements of the form $X^*(M)$ with $m = (0, \ldots, 0, m_{r+1}, \ldots, m_n)$. Then B is a subalgebra of $U(\mathcal{A})$.

Definition 3.205.
Two Lie groups G and G' are ***isomorphic*** *when there exists a differentiable group isomorphism between G and G'.*

They are ***locally isomorphic*** *when there exists neighbourhoods $\mathcal{U}$ and $\mathcal{U}'$ of e and e' and a differentiable diffeomorphism $f : \mathcal{U} \to \mathcal{U}'$ such that*
$\forall x, y, xy \in \mathcal{U},\ f(xy) = f(x)f(y),$
and
$\forall x', y', x'y' \in \mathcal{U}',\ f^{-1}(x'y') = f^{-1}(x')f^{-1}(y').$

Now a great theorem without proof:

Theorem 3.206.
Two Lie groups are locally isomorphic if and only if their Lie algebras are isomorphic.

The following universal property of the *universal* enveloping algebra explains the denomination:

Proposition 3.207.
Let $\sigma : \mathcal{G} \to \mathcal{U}(\mathcal{G})$ the canonical inclusion and A an unital complex associative algebra. A linear map $\varphi : \mathcal{G} \to A$ such that

$$\varphi[X, Y] = \varphi(X)\varphi(Y) - \varphi(Y)\varphi(X) \tag{3.421}$$

can be extended in only one way to an algebra homomorphism $\varphi_0 : \mathcal{U}(\mathcal{G}) \to A$ such that $\varphi_0 \circ \sigma = \varphi$ and $\varphi(1) = 1$

For a proof, see [6].

3.15.1 Adjoint map in $\mathcal{U}(\mathcal{G})$

We know that $\mathrm{Ad}(g) : \mathcal{G} \to \mathcal{G}$ fulfils

$$\mathrm{Ad}(g)[X, y] = [\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y],$$

[32] À la page /23

and we can define $\mathrm{Ad}(g)\colon \mathcal{G} \to \mathcal{U}(\mathcal{G})$ by $\mathrm{Ad}(g)X = X$ where in the right hand side, X denotes the class of X for the quotients of the tensor algebra which defines the universal enveloping algebra.

When $[A, B]$ is seen in $\mathcal{U}(\mathcal{G})$, we have $[A, B] = A \otimes B - B \otimes A$. Then $\mathrm{Ad}(g)\colon \mathcal{G} \to \mathcal{U}(\mathcal{G})$ fulfils proposition 3.207 and is extended in an unique way to $\mathrm{Ad}(g)\colon \mathcal{U}(\mathcal{G}) \to \mathcal{U}(\mathcal{G})$ with $\mathrm{Ad}(g)1 = 1$.

Lemma 3.208.
If $D \in \mathcal{U}(\mathcal{G})$, the following properties are equivalent:

- $D \in \mathcal{Z}(\mathcal{G})$
- $D \otimes X = X \otimes D$ *for all* $X \in \mathcal{G}$
- $e^{\mathrm{ad}\, X} D = D$ *for all* $X \in \mathcal{G}$
- $\mathrm{Ad}(g)D = D$ *for all* $g \in G$.

3.15.2 Invariant fields

If $X \in \mathfrak{g}$, we have the associated left invariant vector field on G given by $\tilde{X}_x = dL_x X$. That field is left invariant as operator on the functions because

$$\tilde{X}_x(u) = \tilde{X}_e(L_x^* u) \tag{3.422}$$

as the following computation shows

$$\tilde{X}_e(L^* u) = \frac{d}{dt}\Big[(L_x^* u)\big(e^{tX}\big)\Big]_{t=0} = \frac{d}{dt}\Big[u\big(xe^{tX}\big)\Big]_{t=0} = \frac{d}{dt}\Big[u\big(\tilde{X}_x(t)\big)\Big]_{t=0} = \tilde{X}_x(u) \tag{3.423}$$

because the path defining $\tilde{X}_x$ is xe^{tX}.

We can perform the same construction in order to build left invariant fields based on $\mathcal{U}(\mathfrak{g})$. If X and Y are elements of $\mathfrak{g}$, the differential operator on $C^\infty(G)$ associated to $XY \in \mathcal{U}(\mathfrak{g})$ is given by

$$(XY)(f) = \frac{d}{dt}\frac{d}{ds}\Big[f\big(X(s)Y(t)\big)\Big]_{\substack{s=0\\t=0}} \tag{3.424}$$

The path defining the field $\widetilde{XY}$ is

$$\widetilde{XY}_x = xX(s)Y(t). \tag{3.425}$$

Thus we have

$$\widetilde{(XY)}_e(L^* u) = \widetilde{(XY)}_x u \tag{3.426}$$

Lemma 3.209.
If $X, Y \in \mathfrak{g}$ we have

$$[\mathrm{ad}(X), \mathrm{ad}(Y)] = \mathrm{ad}([X, Y]). \tag{3.427}$$

Proof. Let $f \in \mathfrak{g}$ and compute the action of $[\mathrm{ad}(X), \mathrm{ad}(Y)]$:

$$\begin{aligned}
[\mathrm{ad}(X), \mathrm{ad}(Y)]f &= \mathrm{ad}(X)[Yf, fY] - \mathrm{ad}(Y)(Xf - fX) && (3.428a)\\
&= (XY - YX)f + f(YX - XY) && (3.428b)\\
&= \mathrm{ad}([X, Y])f. && (3.428c)
\end{aligned}$$

□

3.15.3 Representation of Lie groups

Proposition 3.210.
Let G be a Lie group and $\mathcal{G}$ its Lie algebra. A representation $\varphi\colon G \to \mathrm{End}(V)$ of the group induces a representation $\phi\colon \mathcal{U}(\mathcal{G}) \to \mathrm{End}(V)$ of the universal enveloping algebra with the definitions

$$\begin{aligned}
\phi(X) &= d\varphi_e(X), && (3.429a)\\
\phi(XY) &= \phi(X) \circ \phi(Y) && (3.429b)
\end{aligned}$$

where e is the unit in G and X, Y are any elements of $\mathcal{G}$.

Proof. We have

$$\phi(X) = \frac{d}{dt}\Big[\varphi(e^{tX})v\Big]_{t=0} = d\varphi_e(X)v. \tag{3.430}$$

Notice that, by linearity of the action of $\varphi(e^{tX})$ on v, one can leave v outside the derivation. Now, neglecting the second order terms in t in the derivative, and using the Leibnitz formula, we have

$$\begin{aligned}
\phi([X,Y])v &= \frac{d}{dt}\Big[\varphi(e^{tXY}e^{-tXY})\Big]_{t=0}v \\
&= \frac{d}{dt}\Big[\varphi(e^{tXY})\varphi(\mathbb{1})\Big]_{t=0}v + \frac{d}{dt}\Big[\varphi(\mathbb{1})\varphi(e^{-tXY})\Big]_{t=0}v \\
&= \phi(XY)v - \phi(YX)v \\
&= \big(\phi(X)\phi(Y) - \phi(Y)\phi(X)\big)v \\
&= [\phi(X), \phi(Y)]v,
\end{aligned} \tag{3.431}$$

which is the claim. □

3.16 Representations

References for Lie algebras and their modules are [7–9, 21, 22, 24, 26].

Since $\mathfrak{h}$ is abelian, the operators H_{α_j} $(j = 1, \ldots, l)$ are simultaneously diagonalisable. In that basis of the representation space W, the basis vectors are denoted by $|u_\Lambda\rangle$ and have the property

$$H_{\alpha_i}|u_\Lambda\rangle = \Lambda(H_{\alpha_i})|u_\Lambda\rangle, \tag{3.432}$$

and, as notation, we note $\Lambda_i = \Lambda(H_{\alpha_i})$. The root Λ is a **weight** of the vector $|u_\Lambda\rangle$. The vector $E_\beta|u_\Lambda\rangle$ is of weight $\beta + \Lambda$, indeed,

$$H_{\alpha_i}E_\beta|u_\Lambda\rangle = \big([H_{\alpha_i}, E_\beta] + E_\beta H_{\alpha_i}\big)|u_\Lambda\rangle = \left(\frac{2(\alpha_i,\beta)}{(\alpha_i,\alpha_i)} + \Lambda_i\right)E_\beta|u_\Lambda\rangle. \tag{3.433}$$

Thus the eigenvalue of $E_\beta|u_\Lambda\rangle$ for H_{α_i} is, according to the relation, (3.357), $\beta(H_{\alpha_i}) + \Lambda(H_{\alpha_i})$.

We suppose that the roots α_i are given in increasing order:

$$\alpha_1 \geqslant \alpha_2 \geqslant \ldots \geqslant \alpha_l, \tag{3.434}$$

and one says that a weight is **positive** if its first non vanishing component is positive. Then one choose a basis of W

$$|u_{\Lambda^{(1)}}\rangle, \ldots, |u_{\Lambda^{(N)}}\rangle \tag{3.435}$$

of weight vectors. One say that this basis is **canonical** if

$$\Lambda^{(1)} \geqslant \ldots \geqslant \Lambda^{(N)}. \tag{3.436}$$

Theorem 3.211.
A vector if weight Λ which is a combination of vectors of weight $\Lambda^{(k)}$ all different of Λ vanishes.

Proof. No proof. □

A consequence of that theorem is that, if W is a representation of dimension N of $\mathfrak{g}$, there are at most N different weights. When several vectors have the same weight, the number of linearly independent such vectors is the **multiplicity** of the weight. A weight who has only one weight vector is **simple**.

Proposition 3.212.
The weights Λ and $\Lambda - 2\alpha(\Lambda,\alpha)/(\alpha,\alpha)$ have the same multiplicity for every root α.

Theorem 3.213.
Two representation are equivalent when they have the same highest weight.

Proposition 3.214.
For any weight M and root α,

$$\frac{2(M,\alpha)}{(\alpha,\alpha)} \in \mathbb{Z}, \tag{3.437}$$

and

$$M - \frac{2(M,\alpha)}{(\alpha,\alpha)}\alpha \tag{3.438}$$

is a weight.

Notice, in particular, that for every weight M, the root $-M$ is also a weight.

3.16.1 About group representations

Let π be a representation of a group G. The **character** of π is the function

$$\begin{aligned} \chi_\pi \colon G &\to \mathbb{C} \\ g &\mapsto \operatorname{Tr}\big(\pi(g)\big). \end{aligned} \tag{3.439}$$

From the cyclic invariance of trace, it fulfils $\chi_\pi(gxg^{-1}) = \chi_\pi(x)$, so that the character is a central function.

Let G be a Lie group with Lie algebra $\mathfrak{g}$. We denote by $Z_\pm$ the subgroup of G generated by $\mathfrak{n}^\pm$. The **Cartan subgroup** D of G is the maximal abelian subgroup of G which has $\mathfrak{h}$ as Lie algebra.

A **character** of an abelian group is a representation of dimension one.

Let T be a representation of G on a complex vector space V. One say that $\xi \in V$ is a **highest weight** if

- $T(z)\xi = \xi$ for every $z \in Z_+$,
- $T(g)\xi = \alpha(g)\xi$ for every $g \in D$.

The function $\alpha \colon D \to \mathbb{C}$ is the **highest weight** of the representation T.

Lemma 3.215.
The function α is a character of the group D.

Proof. The number $\alpha(gg')$ is defined by $T(gg')\xi = \alpha(gg')\xi$. Using the fact that T is a representation, one easily obtains $T(gg')\xi = \alpha(g)\alpha(g')\xi$. □

3.16.2 Modules and reducibility

As far as terminology is concerned, one can sometimes find the following definitions. A $\mathfrak{g}$-module is **simple** when the only submodules are $\mathfrak{g}$ and 0. It is **semisimple** when it is isomorphic to a direct sum of simple modules. The module is **indecomposable** if it is not isomorphic to the direct sum of two non trivial submodules.

An vector space endomorphism $a \colon V \to V$ is **semisimple** if V is semisimple as module for the associative algebra spanned by A. In this text, we will not use this terminology but the one in terms of reducibility. It is clear that $\mathfrak{g}$ is itself a $\mathfrak{g}$-module for the adjoint representation. From this point of view, a $\mathfrak{g}$-submodule is an ideal. Then a simple Lie algebra is an irreducible $\mathfrak{g}$-module and a semisimple Lie algebra is completely reducible by corollary 3.46. This explains the terminology correspondence

simple	$\leftrightarrow$	*irreducible*
semisimple	$\leftrightarrow$	*completely reducible.*

Theorem 3.216 (Weyl's theorem)**.**
A representation of a semisimple Lie algebra is completely reducible.

3.16.3 Weight and dual spaces

In general, when $T \colon V \to V$ is an endomorphism of the vector space V and $\lambda \in \mathbb{K}$ ($\mathbb{K}$ is the base field of V), we define

$$V_\lambda = \{v \in V \text{ st } (T - \lambda\mathbb{1})^n v = 0 \text{ for a } n \in \mathbb{N}\}). \tag{3.440}$$

If $V(\lambda) \neq 0$, we say that λ is a **weight** and $V(\lambda)$ is a weight space.

Let now particularize to the case where $\mathfrak{g}$ is a Lie algebra, and $\mathfrak{g}^*$ its dual space (the space of all the complex linear forms on $\mathfrak{g}$). Let ρ be a representation of $\mathfrak{g}$ on a complex vector space V (seen as a $\mathfrak{g}$-module), and $\gamma \in \mathfrak{g}^*$. For each $x \in \mathfrak{g}$, we have $\rho(x) \colon V \to V$ and $\gamma(x) \in \mathbb{C}$; then it makes sense to speak about the operator $\rho(x) - \gamma(x) \colon V \to V$ and to define

$$V_\gamma = \{v \in V \text{ st } \forall x \in \mathfrak{g}, \exists n \in \mathbb{N} \text{ st } \big(\rho(x) - \gamma(x)\big)^n v = 0\}. \tag{3.441}$$

If $V_\gamma \neq 0$, we say that γ is a **weight** for the representation ρ while V_γ is the corresponding **weight space**. A simpler form for complex semisimple Lie algebras will be given in equation (3.105) as corollary of theorem 3.151.

Notice that a root is a weight space for the adjoint representation, see definition 3.76. We denote by Φ the set of non empty root spaces.

Lemma 3.217.
Let $\operatorname{End}(V)$ *be the algebra of linear endomorphism of a vector space* V*. Let* $x_1, \ldots, x_k, y_1, \ldots, y_k \in \operatorname{End}(V)$ *and*

$$e = \sum_i [x_i, y_i].$$

If e *commutes with all* x_i*, then it is nilpotent.*

A proof of this lemma can be found in [17]

Theorem 3.218.
Let $\mathfrak{g}$ *be a Lie algebra of linear endomorphisms of a finite dimensional vector space* V*. We suppose that* V *is a completely reducible* $\mathfrak{g}$*-module and we denote the center of* $\mathfrak{g}$ *by* $\mathcal{Z}$*. Then*

(i) $[\mathfrak{g}, \mathfrak{g}] \cap \mathcal{Z} = 0$*,*

(ii) $L/\mathcal{Z}$ *has a non zero abelian ideal,*

(iii) any element of $\mathcal{Z}$ *is a semisimple endomorphism.*

Problem and misunderstanding 18.
The following proof seems me to be quite wrong.

Proof. Let A be the associative algebra spanned by $\mathfrak{g}$ and the identity on V. It is clear that the A-stable subspaces are exactly the $\mathfrak{g}$-stable ones. Then V is a completely reducible A-module and it has no non zero nilpotent left ideal. Indeed let B be a left ideal in A such that $BB = 0$. In this case, $B \cdot V$ is a A-submodule of V (because B is an ideal) and $V = B \cdot V \oplus W$ for a certain A-submodule W. Since $B \cdot V$ is a A-submodule,

$$B \cdot W \subset (B \cdot V) \cap W$$

(because W is stable under A) which implies $B \cdot W = 0$ and $B \cdot V = B \cdot (BV + W)0$. Consequently, $B = 0$.

Let T be the center of A; this is an ideal, so that T has no non zero nilpotent elements. To see it, consider a nilpotent element $z \in T$. Remark that $T = Az$ is a nilpotent ideal because $AzAz = Az^2A$. Now, we prove that z is a semisimple linear endomorphism of V. By lemma 3.195, with large enough n, $z^n(V)$ finish to stabilize. Let $q = \sum_{v \in V}$ The space $V_s = z^n(V)$ is not zero because z is not nilpotent. Let W be the set of elements of V which are annihilated by a certain power of z. Equation (3.404) makes z semisimple because V_s and W are z-stables.

By lemma 3.217, any element of $[A, A] \cap T$ is nilpotent; but we just saw that it has no non zero nilpotent elements then $[A, A] \cap T = 0$, so that

$$[\mathfrak{g}, \mathfrak{g}] \cap \mathcal{Z} = 0.$$

This proves the first point.

Now we consider an ideal J such that $[J, J] \subset \mathcal{Z}$. Then $[J, J] = [J, J] \cap \mathcal{Z} = 0$. We looks at the abelian ideal $[\mathfrak{g}, J]$ of $\mathfrak{g}$. This is an ideal because $[[g, j], h] = -[[j, h], g] - [[h, g], j]$. By the lemma, the elements of $[\mathfrak{g}, J]$ are nilpotent and the associative algebra generated by $[G, J]$ is also nilpotent because $[\mathfrak{g}, J]$ is abelian.

The elements of B are polynomials with respect to elements of $[\mathfrak{g}, J]$, then $AB \subset BA + B$ because AB is made up with elements of the form $a(hj - jh)^n$ which itself is made up with elements $ah^k j^l$. By commutating j^l, we get

$$j^l a h^k + \text{elements of } [\mathfrak{g}, J],$$

but J is an ideal and $jl \in J$. By induction,

$$(AB)^k \subset B^k A + B^k. \tag{3.442}$$

Since B is nilpotent, AB is a nilpotent left ideal. Then $AB = 0$ which in turn implies $B = 0$. In particular $[\mathfrak{g}, J] = 0$, so that $J \subset \mathcal{Z}$. But any abelian ideal in $\mathfrak{g}/\mathcal{Z}$ is the canonical projection of an ideal J of $\mathfrak{g}$ such that $[J, J] \in \mathcal{Z}$. We conclude that $\mathfrak{g}/\mathcal{Z}$ has no non zero abelian ideal.

□

Now we are able to prove a third version of Lie's theorem:

Theorem 3.219 (Lie).
If $\mathfrak{g}$ *is a solvable ideal, then any completely reducible* $\mathfrak{g}$*-module is annihilated by* $[\mathfrak{g}, \mathfrak{g}]$*.*

Proof. Let V be such a $\mathfrak{g}$-module, ρ the representation of $\mathfrak{g}$ on V and $\mathcal{A} = \rho(\mathfrak{g}) \subset \operatorname{End}(V)$. By assumption, $\mathfrak{a}$ is a solvable subalgebra of $\operatorname{End}(V)$; let $\mathcal{Z}$ be the center of $\mathfrak{a}$. It is clear that $\mathfrak{a}/\mathcal{Z}$ is solvable, so that it has no non zero abelian ideal. But the fact that $\mathfrak{a}/\mathcal{Z}$ is solvable makes one of the $\mathscr{D}^k(\mathfrak{a}/\mathcal{Z})$ an abelian ideal. The conclusion is that $\mathfrak{a}/\mathcal{Z} = 0$, or $\mathfrak{a} = \mathcal{Z}$. Clearly this makes $[\mathfrak{a}, \mathfrak{a}] = 0$. □

Proposition 3.220.
Let $\mathfrak{g}$ be a nilpotent complex Lie algebra and ρ, a representation of $\mathfrak{g}$ on a finite dimensional vector space V. Then

(i) $\forall \gamma \in \mathfrak{g}^*$, *the space V_γ is a $\mathfrak{g}$-submodule of V,*

(ii) if γ is a weight, then there exists a nonzero vector $v \in V_\gamma$ such that $\forall x \in \mathfrak{g}$, $x \cdot v = \gamma(x) v$,

(iii) $V = \bigoplus_\gamma V_\gamma$ where the sum is taken over the set of weight.

From the third point, an element $y \in \mathfrak{g}$ can be decomposed as

$$y = \sum_{\beta \in \Phi} y_\beta \tag{3.443}$$

with $y_\beta \in \mathfrak{g}_\beta$.

From now, we only consider *complex* Lie algebras. A typical milpotent algebra is a Cartan subalgebra of a semisimple Lie algebra.

Proof. Since ρ is a representation,

$$\big(\rho(x) - \gamma(x)\big)\rho(y) = \rho(y)\big(\rho(x) - \gamma(x)\big) + \rho([x, y]).$$

Now let us suppose that $\big(\rho(x) - \gamma(x)\big)^m \rho(y)$ is a sum of endomorphism of the form

$$\rho((\operatorname{ad} x)^p y)\big(\rho(x) - \gamma(x)\big)^q$$

with $p + q = m$. We just saw that it was true for $m = 1$. Let us check for $m + 1$:

$$\begin{aligned} \rho(x)\rho((\operatorname{ad} x)^p y)\big(\rho(x) - \gamma(x)\big)^q &= \rho([x, (\operatorname{ad} x)^p y])\big(\rho(x) - \gamma(x)\big)^q \\ &\quad + \rho((\operatorname{ad} x)^p y)\rho(x)\big(\rho(x) - \gamma(x)\big)^q. \end{aligned} \tag{3.444}$$

Then, since $\mathfrak{g}$ is nilpotent, the space V_γ is a submodule of V because for large enough m and for all y, $\big(\rho(x) - \gamma(x)\big)^m \rho(y) v = 0$ if $v \in V_\gamma$.

Any nilpotent algebra is solvable, then from Lie theorem 3.219, the restrictions of $\rho(x)$ (with $x \in \mathfrak{g}$) to irreducible submodules commute. By Schur's lemma 3.191, they are multiples of identity. But if all $\mathfrak{g}$ is the identity on an irreducible module, then the module has dimension one. In particular, *any irreducible submodule of V_γ has dimension one*[33].

Then, in the weight space V_γ, there is a v which fulfils $\rho(x)v = \lambda(x)v$ for all $x \in \mathfrak{g}$. It is rather clear that it will only works for $\lambda = \gamma$. Our conclusion is that there exists a $v \in V_\gamma$ such that $\rho(x)v = \gamma(x)v$.

Now we consider $\gamma_1, \ldots, \gamma_k$, distinct weights. They are linear forms; then there exists a $x \in \mathfrak{g}$ such that $\gamma_1(x), \ldots, \gamma_k(x)$ are distinct numbers. In fact, the set $\{h \in \mathfrak{h} \text{ st } \alpha_i(h) = \alpha_j(h) \text{ for a certain pair } (i, j)\}$ is a finite union of hyperplanes in $\mathfrak{h}$; then the complementary is non empty.

With this fact we can see that the sum $V_{\gamma_1} + \ldots + V_{\gamma_k}$ is direct. Indeed let $v \in V_{\gamma_i} \cap V_{\gamma_j}$; for the chosen $x \in \mathfrak{g}$ and for suitable m,

$$\big(\rho(x) - \gamma_i(x)\big)^m v = \big(\rho(x) - \gamma_j(x)\big)^m v = 0 \tag{3.445}$$

which implies $\gamma_i(x) = \gamma_j(x)$ or $v = 0$. In particular one has only a finitely many roots and we can suppose that our choice of γ_i is complete.

For $a \in \mathbb{C}$, we define V_a as the set of elements in V which are annihilated by some power of $\rho(x) - a$ with our famous x. By the first lines of the proof, V_a is a $\mathfrak{g}$-submodule of V.

For the same reasons as before[34], if $V_a \neq 0$, there exists a $v \in V_a$ and a weight γ_i such that $\forall y \in \mathfrak{g}$,

$$\rho(y)v = \gamma_i(y)v.$$

[33] Encore que soit pas bien clair pourquoi un tel module existerait... donc l'affiramation suivante ne me semble pas trop justifiée
[34] Celles que je n'ai pas bien comprises

But as v is annihilated by a power of $\big(\rho(x)-a\big)$, it is clear that $a=\gamma_i(x)$, and some theory of linear endomorphism[35] shows that V is the sum of the V_a's:

$$V=\sum_{i=1}^{k}V_{\gamma_i(x)}. \tag{3.446}$$

It remains to be proved that $V_{\gamma_i(x)}\subset V_{\gamma_i}$. Let $y\in\mathfrak{g}$ and

$$V_{i,a}=\{v\in V_{\gamma_i(x)}\text{ st }\exists n:(\rho(y)-a)^n v=0\}.$$

As usual[36] if $V_{i,a}\neq 0$, there exists a $v\in V_{i,a}$ and a weight γ_j such that $\rho(z)v=\gamma_j(z)v$ for any $z\in\mathfrak{g}$. Then $a=\gamma_j(y)=\gamma_i(y)$. But $V_{\gamma_i(x)}$ being the sum of the $V_{i,a}$'s, we have $V_{\gamma_i(x)}=V_{i,\gamma_i(y)}$ for any $y\in\mathfrak{g}$. This makes $V_{\gamma_i(x)\subset V_{\gamma_i}}$.

□

3.16.4 List of the weights of a representation

We consider a representation of highest weight Λ. For each weight M, we define

$$\delta(M)=2\sum_{\alpha_i\in\Pi}M_{\alpha_i} \tag{3.447}$$

where, as usual, $M_\alpha=2(M,\alpha)/(\alpha,\alpha)$. For any root α, we define

$$\gamma(\alpha)=\frac{1}{2}\big(\delta(\Lambda)-\delta(\alpha)\big). \tag{3.448}$$

Proposition 3.214 shows in particular that $\gamma(\alpha)$ is an integer.

Proposition 3.221.
When M is a weight, $\gamma(M)$ is the number of simple roots that have to be subtracted from the highest weight Λ in order to get M.

Proof. No proof. □

Let us consider the sets

$$\Delta_\phi^k=\{M\text{ st }\gamma(M)=k\}. \tag{3.449}$$

That set is the **layer** of order k. Of course, there exists a $T(\phi)$ such that

$$\Delta_\phi=\Delta_\phi^0\cup\Delta_\phi^1\cup\ldots\cup\Delta_\phi^{T(\phi)}. \tag{3.450}$$

That $T(\phi)$ is the **height** of the representation ϕ. If Λ is the highest weight and Λ' is the lowest weight, then we have $\gamma(\Lambda)=0$ and $\gamma(\Lambda')=T(\phi)$.

A corollary of proposition 3.221 is that, if $M\in\Delta_\phi^r$ and if α is a simple root, then $M+\alpha\in\Delta_\phi^{r-1}$, and $M-\alpha\in\Delta_\phi^{r+1}$.

Let us denote by $S_k(\phi)$ the multiplicity of the layer of order k; we have

$$S_0+S_1+\ldots+S_T=N, \tag{3.451}$$

where N is the dimension of the representation ϕ. The number

$$III(\phi)=\max S_k(\phi) \tag{3.452}$$

is the **width** of the representation.

Lemma 3.222.
If Λ is the highest weight and Λ' is the lowest weight, then $\delta(\Lambda)+\delta(\Lambda')=0$.

Proof. No proof. □

From that lemma and the definition of $\gamma(M)$, we deduce that $\delta(\Lambda)-\delta(\Lambda')=2\gamma(\Lambda')=T(\phi)$, so that $\delta(\Lambda)=T(\phi)$ and

$$\delta(M)=T(\phi)-2\gamma(M). \tag{3.453}$$

In particular, $\delta(M)$ has a fixed parity for a given representation ϕ. It is the **parity** (even or odd) of the representation.

[35] théorie que je ne connais pas trop
[36] et comme d'hab, l'argument que je ne saisit pas

Theorem 3.223.
If Λ is the highest weight of the irreducible representation ϕ, then

$$T(\phi) = \sum_{\alpha_i \in \Pi} r_{\alpha_i} \Lambda_\alpha \tag{3.454}$$

where the coefficients r_{α_i} only depend on the algebra, and in particular not on the representation.

Proof. No proof. □

The coefficients r_{α_i} are known for all the simple Lie algebra, see for example page 105 of [21].

3.16.4.1 Finding all the weights of a representation

The following can be found in [21, 22].

Theorem 3.224.
If Δ_ϕ is the weight system of the irreducible representation ϕ, then

$$S_k = S_{T-k} \tag{3.455}$$

and

$$S_r \geqslant S_{r-1} \geqslant \ldots \geqslant S_2 \geqslant S_1 \tag{3.456}$$

where $r = \frac{T}{2} + 1$.

The theorem says that when $T(\phi)$ is even (let us say $T(\phi) = 2r$), then $III(\phi) = S_r(\phi)$ and when $T(\phi)$ is odd (let us say $T(\phi) = 2r + 1$), then

$$III(\phi) = S_r(\phi) = S_{r+1}(\phi). \tag{3.457}$$

Let α be a root. The α**-series** trough the weight M is the sequence of weights

$$M - r\alpha, \ldots, M + q\alpha \tag{3.458}$$

such that $M - (r+1)\alpha$ and $M + (q+1)\alpha$ do not belong to Δ_ϕ.

Proposition 3.225.
Let M be a weight of the representation ϕ and α, any root of $\mathfrak{g}$. If the α-series trough M begins at $M - r\alpha$ and ends at $M + q\alpha$, then

$$\frac{2(M, \alpha)}{(\alpha, \alpha)} = r - q, \tag{3.459}$$

or, more compactly, $M_\alpha + q = r$.

Notice that, in that proposition, q and r are well defined functions of M and α.

We are now able to determine all the weights of the representation ϕ. Let us suppose that we already know all the layers $\Delta^0_\phi, \ldots, \Delta^{r-1}_\phi$. We are going to determine the weights in the layer Δ^r_ϕ.

An element of Δ^r_ϕ has the form $M - \alpha$ with $M \in \Delta^{r-1}_\phi$ and α, a root. Thus, in order to determine Δ^r_ϕ, we have to test if $M - \alpha$ is a weight for each choice of $M \in \Delta^{r-1}_\phi$ and $\alpha \in \Pi$. Using proposition 3.225, if[37]

$$M_\alpha + q \geqslant 1, \tag{3.460}$$

then $M - \alpha \in \Delta_\phi$. The number $M_\alpha - q(M, \alpha)$ is the **lucky number** of the root $M - \alpha$. The root is a weight if its lucky number is bigger or equal to 1. Notice that $q(M, \alpha)$ depends on the representation we are looking at.

Since $M + k\alpha \in \Delta^{r-k}_\phi$, the value of q is known when one knows the "lower" layers. We are thus able to determine, by induction, all the layers from Δ^0_ϕ which only contains the highest weight. For this one, by definition, we always have $q = 0$.

The Dynkin coefficients of one weights can be more easily computed using the following formula, which is a direct consequence of definition of the Cartan matrix:

$$(M - \alpha_j)_i = M_i - A_{ji}. \tag{3.461}$$

As example, let us determine the weights of the representation $\circ\!\!-\!\!\!-\!\!\!-\!\!\overset{1}{\circ}$ of $\mathfrak{su}(3)$. The algebra $\mathfrak{su}(3)$ has two simple roots α and β whose inner products are $(\alpha, \alpha) = (\beta, \beta) = 1$ and $(\alpha, \beta) = -1/2$. The highest weight of $\phi = \circ\!\!-\!\!\!-\!\!\!-\!\!\overset{1}{\circ}$ is $\Lambda = (\alpha + 2\beta)/3$.

[37] At page 104 of [21], that condition is (I think) wrongly written $M_\alpha + q \geqslant 0$; that mistake is repeated in the example of page 106.

We first test if $\Lambda-\alpha$ is a weight. Easy computations show that $\Lambda_\alpha = 0$ wile $q = 0$; thus $\Lambda-\alpha$ is not a weight. The same kind of computations show that $\Lambda_\beta = 1$, so that $\Lambda_\beta = q(\Lambda, \beta) = 1$. That shows that $\Delta^1_\phi = \{\Lambda - \alpha\}$.

Let now $M = \Lambda - \beta = (\alpha - \beta)/3$. Since $M + \alpha \notin \Delta_\phi$, we have $q(M, \alpha) = 0$. On the other hand, $M_\alpha = 1$, so that $M - \alpha \in \Delta^2_\phi$. The last one to have to be tested is $M - \beta$. Since $M + \beta = \Lambda$, we have $q(M, \beta) = 1$, but $M_\beta = -1$. Thus $M_\beta + q(M, \beta) = 0$ and $M - \beta$ is not a weight.

We can obviously continue in that way up to find $\Delta^r_\phi = 0$, but there is an escape to be more rapid. Indeed, using theorem 3.223 with coefficients r_α that can be found in tables (for example in [21]), we find

$$T(\phi) = 2\Lambda_\alpha + 3\Lambda_\beta = 2, \tag{3.462}$$

thus we immediately know that Δ^3_ϕ does not exist.

On the other hand, one knows the width $III(\phi) = \max S_k(\phi)$ because (since $T(\phi) = 2r$, with $r = 1$), we have $III(\phi) = S_1(\phi)$. Thus, once $\Delta^1(\phi)$ is determined, we know that the next ones will never have more elements.

In the example, when we know that $M - \alpha$ is a weight, we do not have to test $M - \beta$.

3.16.5 Tensor product of representations

3.16.5.1 Tensor and weight

Let ϕ and ϕ' be representations of $\mathfrak{g}$ on the vector spaces R and R' of dimensions n and m. If $A \in \mathbb{M}_n(R)$ and $B \in \mathbb{M}_m(R')$, the **tensor product**, also know as the **Kronecker product** of A and B is the matrix $A \otimes B \in \mathbb{M}_{mn}(R \otimes R')$ whose elements are given by

$$C_{ik,jl} = A_{ij} B_{kl}. \tag{3.463}$$

The principal properties of that product are

$$(A_1 A_2) \otimes (B_1 B_2) = (A_1 \otimes B_1)(A_2 \otimes B_2) \tag{3.464a}$$
$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \tag{3.464b}$$
$$\mathbb{1}_R \otimes \mathbb{1}_{R'} = \mathbb{1}_{R \otimes R'} \tag{3.464c}$$

If φ_1 and φ_2 are two representations of a group G, the **tensor product** is defined by

$$(\varphi_1 \otimes \varphi_2)(g) = \varphi_1(g) \otimes \varphi_2(g). \tag{3.465}$$

If ϕ and ϕ' are two representations of a Lie algebra $\mathfrak{g}$, the **tensor product** representation is defined by

$$(\phi \otimes \phi')(X)(v \otimes v') = \big(\phi(X)v\big) \otimes v' + v \otimes \big(\phi'(X)v'\big). \tag{3.466}$$

If $\{\phi_k\}$ are the irreducible representations, a natural question that arise is to determine the coefficients Γ which decompose $\phi \otimes \phi'$ into irreducible representations:

$$\phi \otimes \phi' = \sum_k \Gamma_k(\phi, \phi')\phi_k \tag{3.467}$$

Let W and W' be the representation spaces and consider the following decompositions in weight spaces:

$$W = \bigoplus_{\Lambda \in \Delta_1} W_\Lambda, \qquad W' = \bigoplus_{\Lambda \in \Delta_2} W'_\Lambda. \tag{3.468}$$

By definition,

$$(W \otimes W')_\alpha = \{v \otimes v' \text{ st } (\phi \otimes \phi')(h)(v \otimes v') = \alpha(h)(v \otimes v')\}. \tag{3.469}$$

If $\big(\phi(h)v\big) \otimes v' + v \otimes \big(\phi'(h)v'\big)$ is a multiple of $v \otimes v'$, one requires that

$$\phi(h)v = \alpha_1(h)v, \tag{3.470a}$$
$$\phi'(h)v = \alpha_2(h)v' \tag{3.470b}$$

for the weights α_1 and α_2 of ϕ and ϕ'. Thus we have

$$(W \otimes W')_{\alpha_1 + \alpha_2} = W_{\alpha_1} \otimes W_{\alpha_2}. \tag{3.471}$$

We have in particular that the simple root system $\Delta_{\phi \otimes \phi'}$ of the representation $\phi \otimes \phi'$ is given by

$$\Delta_{\phi \otimes \phi'} = \Delta_\phi + \Delta_{\phi'}. \tag{3.472}$$

What we proved is[38]

[38] The second part is not proved.

Proposition 3.226.
If ϕ is a representation of highest weight Λ and ϕ' is a representation of highest weight Λ', then $\phi \otimes \phi'$ is a representation of height weight $\Lambda + \Lambda'$.

If, moreover, ϕ and ϕ' are irreducible, then $\phi \otimes \phi'$ is irreducible.

An irreducible representation that cannot be written under the form of a tensor product of irreducible representations is a **basic representation**.

Lemma 3.227.
A representation is basic if and only if its highest weight Λ is such that the Λ_{α_i} are all zero but one which is 1.

The basic representations of $\mathfrak{so}(10)$ are given by the Dynkin diagrams of figure 3.1. All the irreducible representations are obtained by tensor products of the basic ones. An **elementary** is a basic representation which has his "1" on a terminal point of the Dynkin diagram.

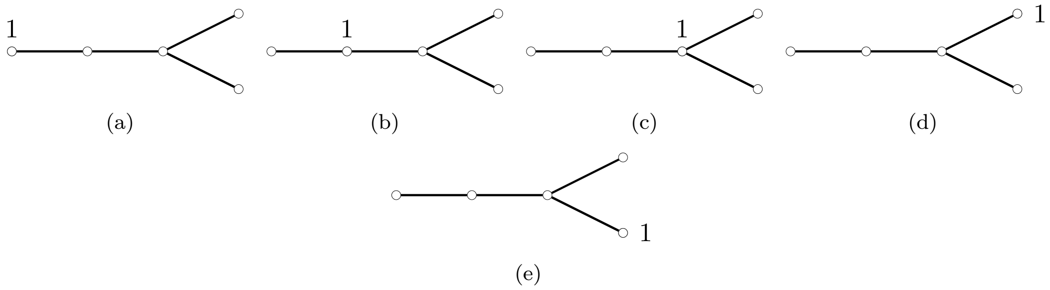


Figure 3.1: Basic representations of $\mathfrak{so}(10)$

3.16.5.2 Decomposition of tensor products of representations

Proposition 3.226 allows us to decompose a tensor product of representations into irreducible representations. Let us do it on a simple example in $\mathfrak{su}(3)$. We consider the representations $\phi = \overset{1}{\circ}\!\!-\!\!-\!\!\circ$ and $\phi' = \circ\!\!-\!\!-\!\!\overset{1}{\circ}$. The first representation has weights

$$\Delta_\phi = \left\{ \frac{\alpha+2\beta}{3}, \frac{\alpha-\beta}{3}, \frac{-(2\alpha+\beta)}{3} \right\}, \tag{3.473}$$

and the second one has

$$\Delta_{\phi'} = \left\{ \frac{\alpha+2\beta}{3}, \frac{\alpha-\beta}{3}, \frac{-(2\alpha+\beta)}{3} \right\}. \tag{3.474}$$

According to equation (3.472), we have 9 weights in the representation $\phi \otimes \phi'$ (all the sums of one element of Δ_ϕ with a one of $\Delta_{\phi'}$). The highest one is

$$\frac{2\alpha+4\beta}{3},$$

which is the double of the highest weight in $\circ\!\!-\!\!-\!\!\overset{1}{\circ}$, so $\phi \otimes \phi'$ contains the representation $\circ\!\!-\!\!-\!\!\overset{2}{\circ}$. Now, we remove from the list of weights of $\phi \otimes \phi'$ the list of weight of $\circ\!\!-\!\!-\!\!\overset{2}{\circ}$; the result is

$$\frac{2\alpha+\beta}{3}, \frac{-(\alpha-\beta)}{3}, \frac{-(\alpha+2\beta)}{3}, \tag{3.475}$$

which are the weights of $\overset{1}{\circ}\!\!-\!\!-\!\!\circ$. The conclusion is that

$$\circ\!\!-\!\!-\!\!\overset{1}{\circ} \otimes \circ\!\!-\!\!-\!\!\overset{1}{\circ} = \circ\!\!-\!\!-\!\!\overset{2}{\circ} \oplus \overset{1}{\circ}\!\!-\!\!-\!\!\circ. \tag{3.476}$$

That procedure of decomposition is quite long because it requires to compute the complete set of weights for some intermediate representations.

3.16.5.3 Symmetrization and anti symmetrization

Let ϕ be a irreducible representation. We want to compute the symmetric and antisymmetric parts of the representation $\phi^{\otimes k} = \underbrace{\phi \otimes \ldots \otimes \phi}_{k \text{ times}}$. These symmetric and antisymmetric parts are denoted by $\phi_s^{\otimes k}$ and $\phi_a^{\otimes k}$ respectively.

Proposition 3.228.
If $\{\xi_1, \ldots, \xi_N\}$ is a canonical basis of ϕ and if we denote by Λ_i the weight of the vector ξ_i, the followings hold:

(i) the weight system of $\phi_a^{\otimes k}$ is

$$\Lambda_{i_1} + \Lambda_{i_2} + \ldots + \Lambda_{i_k} \tag{3.477}$$

with $i_k > \ldots > i_2 > i_1$, and the highest weight is

$$\Lambda_1 + \ldots + \Lambda_k. \tag{3.478}$$

The dimension of the representation $\phi_a^{\otimes k}$ is

$$N\big(\phi_a^{\otimes k}\big) = \binom{n}{k}. \tag{3.479}$$

(ii) The weight system of the representation $\phi_s^{\otimes k}$ is

$$\Lambda_{i_1} + \Lambda_{i_2} + \ldots + \Lambda_{i_k} \tag{3.480}$$

with $i_k \geqslant \ldots \geqslant i_2 \geqslant i_1$, and the highest weight is

$$k\Lambda_1 \tag{3.481}$$

The dimension of the representation $\phi_s^{\otimes k}$ is

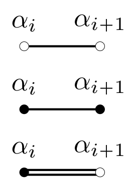
$$N\big(\phi_s^{\otimes k}\big) = \binom{n+k}{k}. \tag{3.482}$$

Proof. No proof. □

The representations $\phi_a^{\otimes k}$ and $\phi_s^{\otimes k}$ might be decomposable and we denote by $\phi_{s>}^{\otimes k}$ and $\phi_{a>}^{\otimes k}$ their highest weight parts.

Let α be a terminal point in a Dynkin diagram. The **branch** of α is the sequence of point of the Dynkin diagram $\alpha = \alpha_1, \alpha_2, \ldots, \alpha_k$ defined by the following properties.

- The point α_i is connected with (and only with) the points α_{i-1} and α_{i+1},
- the connexion between α_i and α_{i+1} is of one of the following forms



(3.483)

- the sequence $\alpha_1, \ldots, \alpha_k$ is maximal in the sense that no α_{k+1} can be added without violating one of the two first rules.

Proposition 3.229.
Let α be a terminal point in a Dynkin diagram and $\alpha_1, \ldots, \alpha_k$ be the corresponding branch. Then we have

$$\phi_{\alpha_r} \simeq \phi_{\alpha\, a>}^{\otimes r} \tag{3.484}$$

for every $r = 1, 2, \ldots, k$.

3.17 Verma module

Let us give the definition of [27]. When $\mathfrak{g}$ is a semisimple Lie algebra, we have the usual decomposition

$$\mathfrak{g} = \mathfrak{n}^- \oplus \mathfrak{h} \oplus \mathfrak{n}^+, \tag{3.485}$$

where each of the three components are Lie algebras. In particular, the universal enveloping algebra $\mathcal{U}(\mathfrak{n}^-)$ makes sense. Let $\mu \in \mathfrak{h}^*$. We build a representation π_μ of $\mathfrak{g}$ on $V_\mu = \mathcal{U}(\mathfrak{n}^-)$ in the following way

- If $Y_\alpha \in \mathfrak{n}^-$, we define

$$\pi_\mu(Y_\alpha)1 = Y_\alpha \tag{3.486a}$$
$$\pi_\mu(Y_{\alpha_1} \ldots Y_{\alpha_n}) = Y_\alpha Y_{\alpha_1} \ldots Y_{\alpha_n}, \tag{3.486b}$$

- if $H \in \mathfrak{h}$, we define

$$\pi_\mu(H)1 = \mu(H) \tag{3.487a}$$
$$\pi_\mu(Y_{\alpha_1} \ldots Y_{\alpha_k}) = \big(\mu(H) - \sum_{j=1}^{k} \alpha_j(H)\big) Y_{\alpha_1} \ldots Y_{\alpha_k}, \tag{3.487b}$$

- and if $X_\alpha \in \mathfrak{n}^+$, we define

$$\pi_\mu(X_\alpha)1 = 0 \tag{3.488a}$$
$$\pi_\mu(X_\alpha) Y_{\alpha_1} \ldots Y_{\alpha_k} = Y_{\alpha_1}\big(\pi_\mu(X_\alpha) Y_{\alpha_2} \ldots Y_{\alpha_k}\big) \tag{3.488b}$$
$$- \delta_{\alpha,\alpha_1} \sum_{j=1}^{k} \alpha_j(H_\alpha) Y_{\alpha_1} \ldots Y_{\alpha_k}. \tag{3.488c}$$

In the last one, we do an inductive definition.

Lemma 3.230.
The couple (π_μ, V_μ) is a representation of $\mathfrak{g}$ on V_μ.

Proof. No proof. □

That representation is one **Verma module** for $\mathfrak{g}$. If the algebra $\mathfrak{g}$ is an algebra over the field $\mathbb{K}$, the field $\mathbb{K}$ itself is part of $\mathcal{U}(\mathfrak{n})^-$, so that the scalars are vectors of the representation. In that context, the multiplicative unit $1 \in \mathbb{K}$ is denoted by v_0.

Theorem 3.231.
The representation (π_μ, V_μ) of the semisimple Lie algebra $\mathfrak{g}$ is a cyclic module of highest weight, with highest weight μ and where v_0 is a vector of weight μ.

Proof. No proof. □

The Verma module is, *a priori*, infinite dimensional and non irreducible, thus one has to perform quotients of the Verma module in order to build finite dimensional irreducible representations.

3.18 Cyclic modules and representations

An example over $\mathfrak{so}(3)$ is given in subsection **??**. The case of $\mathfrak{so}(5)$ is treated in subsection **??**. Let $\mathfrak{g}$ be a semisimple Lie algebra with a Cartan subalgebra $\mathfrak{h}$ and a basis Δ for its roots $\Phi = \Phi^+ \cup \Phi^-$. Let W be a finite dimensional $\mathfrak{g}$-module.

Lemma 3.232.
If $\mathfrak{g}$ is a nilpotent complex algebra and if γ is a weight, then there exists a v in V_γ such that $c \cdot v = \gamma(x)v$ for every $x \in \mathfrak{g}$.

This is the proposition 3.220. Notice that a Cartan algebra is nilpotent, thus one has at least one vector of W which is a common eigenvector of every elements of $\mathfrak{h}$, in other words, $\exists \mu \in \mathfrak{h}^*$ and $\exists w \in W$ such that

$$hw = \mu(h)w \tag{3.489}$$

for every $h \in \mathfrak{h}$, and $w \neq 0$. If w is such and if $x \in \mathfrak{g}_\alpha$, we have

$$(hx) \cdot w = [h, x] \cdot w + (xh) \cdot w = \alpha(h)x \cdot w + x\mu(h)w = (\alpha + \mu)(h)x \cdot w. \tag{3.490}$$

If we define

$$S = \{w \in W \text{ st } \exists \mu \in \mathfrak{h}^* \text{ st } hw = \mu(h)w\}, \tag{3.491}$$

this is not a vector space, but the vector space $\operatorname{Span} S$ generated by S is invariant under $\mathfrak{g}$ because S itself is invariant under all the $\mathfrak{g}_\alpha$ with $\alpha \in \mathfrak{g}^*$.

On the other hand, we suppose that $\mathfrak{g}$ and W are finite dimensional, so that their dual are isomorphic. Since a Cartan subalgebra is chosen, we have the decomposition

$$\mathfrak{g} = \mathfrak{h} \oplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}_\alpha \tag{3.492}$$

where $\mathfrak{g}_\alpha = \{x \in \mathfrak{g} \text{ st } [h, x] = \alpha(h)x \, \forall g \in \mathfrak{h}\}$. When $\alpha \in \mathfrak{h}^*$, the two following spaces are independent of the choice of the Cartan subalgebra $\mathfrak{h}$:

$$\begin{aligned} W_\alpha &= \{v \in W \text{ st } hv = \alpha(h)v \, \forall h \in \mathfrak{h}\} \\ \mathfrak{g}_\alpha &= \{x \in \mathfrak{g} \text{ st } [h, x] = \alpha(h)x \, \forall h \in \mathfrak{h}\}. \end{aligned} \tag{3.493}$$

If $v_\alpha \in W_\alpha$ and $x_\beta \in \mathfrak{g}_\beta$, we have

$$h(x_\beta v_\alpha) = \big([h, x_\beta] + x_\beta h\big)v_\alpha = \big(\beta(h) + \alpha(h)\big)x_\beta v_\alpha, \tag{3.494}$$

so $x_\beta v_\alpha \in W_{\alpha+\beta}$. Thus x_β is a map

$$x_\alpha \colon W_\alpha \to W_{\alpha+\beta}. \tag{3.495}$$

Since W is finite dimensional, there exists a maximal α such that $W_\alpha \neq 0$. We name it λ. For every $\beta \in \Phi^+$, we have $W_{\lambda+\beta} = \{0\}$. In particular, if $v_\lambda \in W_\lambda$,

$$x_\alpha x_\lambda = 0 \tag{3.496}$$

for every $\alpha \in \Phi^+$, and, of course,

$$hv_\lambda = \lambda(h)v_\lambda. \tag{3.497}$$

On the other hand, for every vector $v \in W$, and for v_λ in particular, the space $\mathcal{U}(\mathfrak{g})v$ is invariant, so

$$W = \mathcal{U}(\mathfrak{g})v_\lambda \tag{3.498}$$

by irreducibility. One say that W is the **cyclic module** generated by v_λ.

3.18.1 Choice of basis

Theorem 3.233.
Let $\mathfrak{g}$ be a Lia algebra on a field of characteristic zero. If $\{x_i\}$ is an ordered basis of $\mathfrak{g}$, then

$$\{x_{i_1} \cdots x_{i_n} \text{ st } i_1 \leqslant \ldots \leqslant i_n\} \tag{3.499}$$

is a basis for the universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ of $\mathfrak{g}$.

One can find a proof in [7].

3.18.2 Roots and highest weight vectors

Proposition 3.234.
An irreducible cyclic module is generated by the elements of the form $f_1^{i_1} \cdots f_m^{i_m} v_\lambda$.

Proof. From theorem 3.233, the monomials of the form

$$(f_1^{i_1} \cdots f_m^{i_m}) \cdot (h_1^{j_1} \cdots h_l^{j_l}) \cdot (e_1^{k_1} \cdots e_m^{k_m}) \tag{3.500}$$

form a basis of $\mathcal{U}(\mathfrak{g})$. When one act with such an element on v_λ, the e_i kill it, while the h_i do not act (a part of changing the norm). Thus, in fact, the module W is generated by the only elements $f_1^{i_1} \cdots f_m^{i_m} v_\lambda$ □

In very short, one can write

$$W = (\mathfrak{n}^-)^n v_\lambda. \tag{3.501}$$

Since $f_k v_\alpha \in \mathfrak{g}_{\alpha-\alpha_k}$, we have

$$f_1^{i_1} \cdot f_m^{i_m} v_\lambda \in \mathfrak{g}_{\lambda-(i_m\alpha_m-\ldots i_1\alpha_1)}. \tag{3.502}$$

The set of roots is ordered by

$$\mu_1 < \mu_2 \quad \text{iff} \quad \mu_2 - \mu_1 = \sum_i k_i\alpha_i \tag{3.503}$$

with $\alpha_i > 0$ and with $k_i \in \mathbb{N}$. Equation (3.502) means that

$$\mu < \lambda \tag{3.504}$$

for every weight μ of W.

Definition 3.235.
*Let $\mathfrak{g}$ be a finite dimensional Lia algebra. A **cyclic module of highest weight** for $\mathfrak{g}$ is a module (not specially of finite dimension) in which there exists a vector v_+ such that $x_+v_+ = 0$ for every $x_+ \in \mathfrak{n}^+$ and $hv_+ = \lambda(h)v_+$ for every $h \in \mathfrak{h}$.*

Proposition 3.236.
Every submodule of a cyclic highest weight module is a direct sum of weight spaces.

Proof. No proof. □

From the relation $x_+v_+ = 0$, we know that all the weight spaces satisfy V_μ satisfy $\mu < \lambda$, and, since a module is the sum of all its submodules,

$$V = \bigoplus V_\mu. \tag{3.505}$$

Notice that if v_+ is in a submodule, then that submodule is the whole V, thus the sum of two proper submodules is a proper submodule. We conclude that V has an unique maximal submodule, and has thus an unique irreducible quotient.

3.18.3 Dominant weight

We know that every representation is defined by a highest weight. The following proposition[28] shows that every root cannot be a highest weight of an irreducible representation.

Proposition 3.237.
The highest weight of an irreducible representation of a simple complex Lie algebra is an integral dominant weight.

Proof. Let α_i be a simple root and consider the corresponding copy of $\mathfrak{sl}(2,\mathbb{C})$ generated by $\{e_i, f_i, h_i\}$ (see subsection 3.8.4). The following part of $L(\Lambda)$ is a $\mathfrak{sl}(2,\mathbb{C})_i$-module:

$$V(\alpha_i) = \bigoplus_{n\in\mathbb{Z}} V_{\Lambda+n\alpha_i} = V_\Lambda \oplus V_{\Lambda-\alpha_i} \oplus V_{\Lambda-2\alpha_i} \oplus \ldots \oplus V_{\Lambda-r\alpha_i} \tag{3.506}$$

for some positive integer r. Notice that the sum over $n \in \mathbb{Z}$ does not contain terms with $n < 0$ because Λ being an highest weight, $V_{\Lambda+k\alpha_i} = \varnothing$ when $k > 0$. We know that in a $\mathfrak{sl}(2,\mathbb{C})$-module the eigenvalues of h run from $-m$ to m (see equations (**??**) for example). Thus here

$$\Lambda(h_i) = -(\Lambda - r\alpha_i)(h_i). \tag{3.507}$$

By construction $\alpha_i(h_i) = 2$, so $\Lambda(h_i) = r$ and the proof is finished. □

Proposition 3.238.
If Λ is the highest weight of the representation $L(\Lambda)$ of the complex simple Lie algebra $\mathfrak{g}$ and if w_0 is the longest elements of the Weyl group, then $w_0\Lambda$ is the lowest weight.

Proof. First remember that whenever λ is a weight of a representation and w is an element of the Weyl group, the root $w\lambda$ is a weight[39]; in particular $w_0\Lambda$ is a weight of $L(\Lambda)$. Let $v \in L(\Lambda)_{w_0\Lambda}$; we want to show that $X_i^- v = 0$.

If $X_i^- v \neq 0$, then $w_0\Lambda - \alpha_i$ is a weight and $w_0(w_0\Lambda - \alpha_i) = \Lambda - w_0\alpha_i$ is a weight too. Here we used the fact that $w_0^2 = \text{id}$. □

[39] To be proved.

Problem and misunderstanding 19.
Still to be shown:

(i) $w\lambda$ *is a weight*

(ii) $w_0^2 = \mathrm{id}$

3.18.4 Verma modules

Let us consider

$$\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}^+, \tag{3.508}$$

and take $\alpha \in \mathfrak{h}^*$. Now, we define $\mathbb{C}_\alpha$ as the vector space $\mathbb{C}$ (one dimensional, generated by $z_+ \in \mathbb{C}$) equipped with the following action of $\mathfrak{b}$:

$$\big(h + \sum_{\mu<0} x_\mu\big) z_+ = \alpha(h) z_+. \tag{3.509}$$

The vector space $\mathbb{C}_\alpha$ becomes a left $\mathcal{U}(\mathfrak{b})$-module. On the other hand, $\mathcal{U}(\mathfrak{g})$ is a free right $\mathcal{U}(\mathfrak{b})$-module because $\mathcal{U}(\mathfrak{b}) \cup \mathcal{U}(\mathfrak{g}) \subseteq \mathcal{U}(\mathfrak{g})$. As $\mathcal{U}(\mathfrak{b})$-module, a basis of $\mathcal{U}(\mathfrak{g})$ is given by $\mathfrak{n}^-$, i.e. by $\{f_1^{i_1} \cdots f_m^{i_l}\}$. The **Verma module** is the cyclic module

$$\mathrm{Verm}(\alpha) = \mathcal{U}(\mathfrak{g}) \otimes_{\mathcal{U}(\mathfrak{b})} \mathbb{C}_\alpha \tag{3.510}$$

which has a highest weight vector $v_\lambda = 1 \otimes z_+$. The tensor product over $\mathcal{U}(\mathfrak{b})$ beans that, when $X \in \mathcal{U}(\mathfrak{g})$, then

$$\big(h + \sum_\mu x_\mu\big) X \otimes_{\mathcal{U}(\mathfrak{b})} z z_+ = X \otimes \big(h + \sum_\mu x_\mu\big) z z_+ = X \otimes_{\mathcal{U}(\mathfrak{g})} z\alpha(h) z_+ = \alpha(h) X \otimes_{\mathcal{U}(\mathfrak{b})} z z_+. \tag{3.511}$$

The Verma module is generated by $1 \otimes z_+$ and the fact that

$$zX(1 \otimes z_+) = X \otimes z z_+. \tag{3.512}$$

Proposition 3.239.
Two irreducible cyclic modules with same highest weight are isomorphic.

Proof. Let V and W be two highest weight cyclic modules with highest weight λ and highest weight vectors v_λ and w_λ. In the module $V \oplus W$, the vector $v_\lambda \oplus w_\lambda$ is a highest weight vector of weight λ. Let us consider the module

$$Z = \mathcal{U}(\mathfrak{g})(v_\lambda \oplus w_\lambda). \tag{3.513}$$

That module is a highest weight cyclic module. The projections onto $V = Z/W$ and $W = Z/V$ are non vanishing surjective homomorphisms, so V and W are irreducible quotients of Z. But we saw bellow equation (3.505) that Z can only accept one irreducible quotient. Thus V and W are isomorphic. □

We denote by $\mathrm{Irr}_{\mathfrak{g}}(\lambda)$ the unique cyclic highest weight $\mathfrak{g}$-module with highest weight λ.

3.19 Semi-direct product

3.19.1 From Lie algebra point of view

Here, the matter comes from [13, 29]. When $\mathfrak{a}$ and $\mathfrak{b}$ are Lie algebras, one can consider $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ as vector space, and define a Lie algebra structure on $\mathfrak{g}$ by

$$[(a,b),(a',b')] = ([a,a'],[b,b']).$$

This is the **direct sum** of $\mathfrak{a}$ and $\mathfrak{b}$.

An endomorphism $\mathcal{D}$ of the Lie algebra $\mathfrak{a}$ is a **derivation** when

$$\mathcal{D}[X,Y] = [\mathcal{D}X, Y] + [X, \mathcal{D}Y].$$

The set of the derivations of $\mathfrak{a}$ is written $\mathtt{Der}\,\mathfrak{a}$.

Proposition 3.240.
Let $\mathfrak{a}$ be a Lie algebra

(i) $\mathtt{Der}\,\mathfrak{a}$ *is a Lie algebra for the usual commutator,*

(ii) $\mathrm{ad}\colon \mathfrak{a} \to \mathtt{Der}\,\mathfrak{a} \subseteq \mathrm{End}\,\mathfrak{a}$ *is a Lie algebra homomorphism.*

Proof. For the first statement, we just have to compute to see that if $\mathcal{D}, \mathcal{E} \in \mathtt{Der}\,\mathfrak{a}$,

$$[\mathcal{D}, \mathcal{E}][X, Y] = (\mathcal{D}\mathcal{E} - \mathcal{E}\mathcal{D})[X, Y] = [[\mathcal{D}, \mathcal{E}]X, Y] + [X, [\mathcal{D}, \mathcal{E}]Y].$$

The second comes from the fact that $\operatorname{ad} X \in \mathtt{Der}\,\mathfrak{a}$ for any $X \in \mathfrak{a}$ and $\operatorname{ad}[X, Y] = \operatorname{ad} X \operatorname{ad} Y - \operatorname{ad} Y \operatorname{ad} X$. □

Let us now consider the vector space direct sum $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$. Let us suppose moreover that $\mathfrak{a}$ is a Lie subalgebra of $\mathfrak{g}$ and that $\mathfrak{b}$ is an idea in $\mathfrak{g}$. So we have that

$$\operatorname{ad}|_{\mathfrak{b}} \in \mathtt{Der}\,\mathfrak{b}.$$

By proposition 3.240, we have a homomorphism $\pi\colon \mathfrak{a} \to \mathtt{Der}\,\mathfrak{b}$, $\pi(A) = \operatorname{ad} A|_{\mathfrak{b}}$. So if $A \in \mathfrak{a}$ and $B \in \mathfrak{b}$, $[A, B] = \pi(A)B$. The conclusion is that the Lie algebra structure of $\mathfrak{g}$ is given by $\mathfrak{a}$, $\mathfrak{b}$ and π. In this case, we write $\mathfrak{g} = \mathfrak{a} \oplus_\pi \mathfrak{b}$, and we say that $\mathfrak{g}$ is the semidirect product of $\mathfrak{a}$ and $\mathfrak{b}$. The following theorem gives the general definition of semidirect product.

Theorem 3.241.
Let $\mathfrak{a}$ and $\mathfrak{b}$ be two Lie algebras, and $\pi\colon \mathfrak{a} \to \mathtt{Der}\,\mathfrak{b}$, a Lie algebra homomorphism. There exists an unique Lie algebra structure on the vector space $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ such that

- *the commutators on $\mathfrak{a}$ and $\mathfrak{b}$ are the old ones,*
- *$[A, B] = \pi(A)B$ for any $A \in \mathfrak{a}$ and $B \in \mathfrak{b}$.*

In this case, in the so defined Lie algebra $\mathfrak{g}$, $\mathfrak{a}$ is a subalgebra and $\mathfrak{b}$ is an ideal.

The vector space $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ endowed with this Lie algebra structure is the **semidirect product** of $\mathfrak{a}$ and $\mathfrak{b}$, it is denoted by

$$\mathfrak{g} = \mathfrak{a} \oplus_\pi \mathfrak{b}$$

One also often speak about **split extension** of $\mathfrak{a}$ by $\mathfrak{b}$, with the splitting map π.

Proof. The unicity part is clear: the Lie algebra structure is completely defined by the two conditions and the condition of antisymmetry. The matter is just to see that this structure is a Lie algebra structure: we have to check Jacobi. If in $[[X, Y], Z]$, X, Y, Z are all three in $\mathfrak{a}$ or $\mathfrak{b}$, it is trivial. The two other cases are :

- X, $Y \in \mathfrak{a}$ and $Z \in \mathfrak{b}$. In this case, we use $\pi([X, Y]) = \pi(X)\pi(Y) - \pi(Y)\pi(x)$ (because π is a Lie algebra homomorphism) to find
$$[[X, Y], Z] = \pi([X, Y])Z = -[[Y, Z], X] - [[Z, X], Y].$$
- The second case is X, $Y \in \mathfrak{b}$ and $Z \in \mathfrak{a}$. Here, we use the fact that $\pi(Z)$ is a derivation of $\mathfrak{b}$. The computation is also direct.

It is clear that $\mathfrak{b}$ is an ideal because for any $A \in \mathfrak{a}$ and $B \in \mathfrak{b}$, $[B, A] = -[A, B] = -\pi(A)B \in \mathfrak{b}$.

□

The theory of split extension is often used in the following sense. We have a Lie algebra $\mathfrak{g}$ which decomposes (as vector space) into a direct sum $\mathfrak{a} \oplus \mathfrak{b}$. If in $\mathfrak{g}$ the map $a \mapsto \operatorname{ad}(a)$ is an action of $\mathfrak{a}$ on $\mathfrak{b}$, we say that $\mathfrak{g}$ is a split extension

$$\mathfrak{g} = \mathfrak{a} \oplus_{\mathrm{ad}} \mathfrak{b}.$$

This way to use split extensions is used for example in the proof of proposition **??**.

3.19.2 From a Lie group point of view

.

Definition 3.242.
*A subgroup H is **normal** in the group G if for any $g \in G$ and $a \in H$, $gag^{-1} \in H$.*

If G is a group, N a normal subgroup and L a subgroup, we have $LN = NL$ where, by notation, if A and B are subsets of G, $AB = \{xy | x \in A, y \in B\}$.

If N and L are groups, an **extension** of N by G is a short exact sequence

$$e \longrightarrow N \xrightarrow{i} G \xrightarrow{\pi} L \xrightarrow{L} e \qquad (3.514)$$

which means that

(i) i is injective because only e_N is sent to e_G,

(ii) π is surjective because the whole L is sent to e.

One often say that G is an extension of N by L. In the most common case, i is the inclusion, $L = G/N$ and π is the natural projection.

We say that the extension is **split** when there exists a *split homomorphism* $\rho\colon L \to G$ such that $\rho \circ \pi = \mathrm{id}_G$.

Definition 3.243.
*We say that G is the **semidirect product** of N and L when any $g \in G$ can be written in one and only one way as $g = nl$ with $n \in N$ and $l \in L$.*

Definition 3.244.
*A **Lie group homomorphism** between G and G' is a map $u\colon G \to G'$ which is a group homomorphism and a morphism between G and G' as differentiable manifolds.*

Lemma 3.245.
Any continuous (group) homomorphism between two Lie groups is a Lie *group homomorphism.*

We consider G, a connected Lie group; N, a closed normal subgroup; and L, a connected immersed Lie group. Moreover, we suppose that G is semidirect product of N and L.

Proposition 3.246.
The restriction to L of the canonical projection $\pi\colon G \to G/N$ is continuous for the induced topology from G to L.

Proof. The definition of an open set $\mathcal{U}$ in G/N is that $\pi^{-1}(\mathcal{U})$ is open in G. Then it is clear that π is continuous. The matter is to check it for $\pi|_L$. Let $\mathcal{U}$ be a subset of $\pi(L)$. It is unclear that $\pi^{-1}(\mathcal{U}) \subset L$, but it is true that $\pi|_L^{-1}(\mathcal{U}) \subset L$.

As far as the induced topology on L is concerned, $A \subset L$ is open when $A = \mathcal{O} \cap L$ for a certain open set $\mathcal{O}$ in G.

Let $\mathcal{U}$ be an open subset of $\pi|_L(L)$; this is $\pi^{-1}(\mathcal{U})$ is open in G. We have to compare $\pi^{-1}(\mathcal{U})$ and $\pi|_L^{-1}(\mathcal{U})$. Since

$$\pi|_L^{-1}(\mathcal{U}) = \{x \in L | \pi(x) \in \mathcal{U}\},$$

we have $\pi|_L^{-1}(\mathcal{U}) = \pi^{-1}(\mathcal{U}) \cap L$. But $\pi^{-1}(\mathcal{U})$ is open in G, then $\pi^{-1}(\mathcal{U}) \cap L$ is open in L. □

Proposition 3.247.
The group G is the semidirect product of N and L if and only if $G = NL$ and $N \cap L = \{e\}$.

Proof. If G is semidirect product of N and L, $G = NL$ is clear. In this case, if $e \neq z \in N \cap L$, $z = ez = ze$, thus $z \in G$ can be written in two ways as xy with $x \in N$ and $y \in L$.

For the converse, let us consider $n'l' = nl$. Then $x^{-1}x' = yy'^{-1} \in N \cap L = \{e\}$. Thus $x' = x$ and $y' = y$. □

Now, we consider N, a normal subgroup of G. If $\pi\colon G \to G/N$ is the canonical homomorphism, the restriction $\pi|_L\colon L \to G/N$ is an isomorphism. Indeed, on the one hand, this is surjective because $G = NL$ yields $[g] = [nl] = [l] = \pi|_L(l)$. On the other hand, $\pi|_L(l) = \pi|_L(l')$ implies that $l = nl'$ for a certain $n \in N$. Then $ll'^{-1} = n \in N \cap L = \{e\}$. So $n = e$ and $l = l'$.

Remark 3.248.
If N is any normal subgroup of G, there doesn't exist in general any subgroup L of G such that G should be the semidirect product of N and L.

If G is the semidirect product of N and L, for any $y \in L$, $\sigma_y\colon x \to yxy^{-1}$ is an automorphism of N. The point is that $\sigma_y(a) \in N$ for all $a \in N$ because N is a normal subgroup.

It is also clear that $\forall\, u, v \in L$, $\sigma_{uv} = \sigma_u \circ \sigma_v$. Then $\sigma\colon L \to \operatorname{Aut} N$[40] is a homomorphism. Moreover, the data of σ, N and L determines the law in G (provided the fact that the product NL is seen as formal) because any element of G can be written as nl; thus a product GG is $(nl)(n'l') = (n\sigma_y(n'))(ll')$

Proposition 3.249.
Let N and L be two Lie groups and $\sigma\colon L \to \operatorname{Aut} N$ a homomorphism. With the law

$$(x, y)(x', y') = (x\sigma_y(x'), yy'),$$

[40] $\operatorname{Aut} N$ is the set of all the automorphism of N.

the set $S = N \times L$ is a group.

Proof. If e is the neutral of N and e' the one of L, it is cleat that (e, e') is the neutral of S. It is also easy to check that the inverse of (x, y) is $(\sigma_{y^{-1}}(x^{-1}), y^{-1})$. The associativity is just a computation using $\sigma_y(ab) = \sigma_y(a) \circ \sigma_y(b)$ and $\sigma_x \circ \sigma_y = \sigma_{xy}$. □

The set $N \times L$ endowed with this inner product is denoted

$$N \times_\sigma L.$$

Proposition 3.250.
If G is the semidirect product of N and L, then G is isomorphic to $N \times_\sigma L$.

Proof. The isomorphism is $T \colon N \times_\sigma L \to G$, $T(x, y) = xy$. On the one hand, it is bijective because an element of G can be written as nl with $n \in N$ and $l \in L$ in only one way. On the other hand, it is easy to check that $T((x, y)(x', u')) = T(x, y)T(x', y')$. □

One can now give the final definition. Let us consider two connected Lie groups N, L and a Lie group homomorphism $\sigma \colon L \to \operatorname{Aut} N$. By , the map $N \times L \to N$, $(x, y) \to \sigma_y(x)$ is C^∞. So, the group structure on $N \times L$ given by

$$(x, y)(x', y') = (x\sigma_y(x'), yy') \tag{3.515}$$

is compatible with the C^∞ structure of $N \times L$ (seen as a Lie group). The manifold $N \times L$ endowed with the group structure (3.515) is the **semidirect product** on N and L; this is denoted by

$$N \times_\sigma L.$$

3.19.3 Introduction by exact short sequence

3.19.3.1 General setting

Let G_0, G_1 and G_2 be tree connected Lie groups. A **short exact sequence** between them is two group homomorphisms

$$\begin{aligned} \iota &\colon G_0 \to G_1 \\ \pi &\colon G_1 \to G_2 \end{aligned} \tag{3.516}$$

such that $\operatorname{Im}(\iota) = \operatorname{Ker}(\pi)$. In that case, one says that G_1 is an **extension** of G_2 by G_0.

Since the group $\iota(G_0)$ is the kernel of an homomorphism, it is normal and we write $\iota(G_0) \lhd G_1$. Moreover, $\iota(G_0) = \pi^{-1}(e_2)$ and is then closed in G_1. As group, we have

$$G_2 = G_1/\iota(G_0). \tag{3.517}$$

The extension is **split** if there exists a Lie group homomorphism $j \colon G_2 \to G_1$ such that

$$\pi \circ j = \operatorname{id}|_{G_2}. \tag{3.518}$$

This condition imposes j to be injective. In that case we have an action of G_2 on G_0 defined by

$$\begin{aligned} R &\colon G_2 \to \operatorname{Aut}(G_0) \\ R_{g_2}(g_0) &= \iota^{-1}\Big(\mathbf{Ad}\,\big(j(g_2)\big)\iota(g_0)\Big). \end{aligned} \tag{3.519}$$

Notice that $\mathbf{Ad}\,\big(j(g_2)\big)\iota(g_0)$ belongs to $\iota(G_0)$ because the latter is normal.

As manifold we consider

$$G = G_0 \times G_2 \tag{3.520}$$

and we define the multiplication law

$$\begin{aligned} \cdot &\colon G \times G \to G \\ (g_0, g_2) \cdot (g_0', g_2') &= \big(g_0 R_{g_2}(g_0'), g_2 g_2'\big). \end{aligned} \tag{3.521}$$

For associativity we have

$$(g_0, g_2) \cdot \big((g_0', g_2') \cdot (g_0'', g_2'')\big)) = \Big(g_0 R_{g_2}\big(g_0' R_{g_2'}(g_0'')\big), g_2 g_2' g_2''\Big) \tag{3.522}$$

while

$$\big((g_0, g_2) \cdot (g_0', g_2')\big) \cdot (g_0'', g_2'') = \big(g_0 R_{g_2}(g_0') R_{g_2 g_2'}(g_0''), (g_2' g_2'')\big). \tag{3.523}$$

Thus the product is associative if and only if

$$g_0 R_{g_2}\big(g_0' R_{g_2'}(g_0'')\big) = \big(g_0 R_{g_2}(g_0')\big) R_{g_2 g_2'}(g_0''). \tag{3.524}$$

That equality is in fact true because R is a morphism from G_2 to $\operatorname{Aut}(G_0)$, so that $R_{g_2} R_{g_2'} = R_{g_2 g_2'}$.

The neutral in G is (e_0, e_2).

Since $R_{g_2}(g_0)$ is smooth with respect to both variables, the product is smooth. In that way, G becomes a Lie group named the **semi direct product** of G_2 by G_0 and is denoted by

$$G_0 \rtimes_R G_2. \tag{3.525}$$

All the construction is still valid when R is an homomorphism which does not comes from a split extension.

We define the product $G_0 \times G_2 \to G$ by

$$g_0 \cdot g_2 = (g_0, e_2) \cdot (e_0, g_2) \tag{3.526}$$

The diagram

$$\begin{tikzcd}
 & G_1 \arrow[dr, "\pi"] & \\
G_0 \arrow[ur, "\iota"] \arrow[dr, "\mathrm{id} \times \{e\}"'] & & G_2 \\
 & G \arrow[uu, dotted, "\varphi"] \arrow[ur, "\mathrm{pr}_2"'] &
\end{tikzcd} \tag{3.527}$$

suggests us to define the map

$$\begin{aligned} \varphi\colon G_0 \times G_2 &\to G_1 \\ (g_0, g_2) &\mapsto \iota(g_0) j(g_2) \end{aligned} \tag{3.528}$$

This is a Lie group homomorphism because on the one hand

$$\varphi(g_0, g_2) \cdot \varphi(g_0', g_2') = \iota(g_0) j(g_2) \cdot \iota(g_0') j(g_2'), \tag{3.529}$$

while on the other hand

$$\begin{aligned} \varphi\big((g_0, g_2) \cdot (g_0', g_2')\big) &= \varphi\big(g_0 R_{g_2}(g_0'), g_2 g_2'\big) \\ &= \varphi\Big(g_0 \iota^{-1}\big(\mathbf{Ad}(j(g_2))\iota(g_0')\big), g_2 g_2'\Big) \\ &= \iota\Big(g_0 \iota^{-1}\Big(\mathbf{Ad}(j(g_2))\iota(g_0')\Big)\Big) j(g_2 g_2') \\ &= \iota(g_0) j(g_2) \iota(g_0') j(g_2') \end{aligned} \tag{3.530}$$

because ι and j are homomorphisms.

The Leibniz rule on $\iota(g_0) j(g_2)$ provides the differential

$$(d\varphi)_e = (d\iota)_{e_0} \oplus (dj)_{e_2}. \tag{3.531}$$

This is injective because j is injective. The kernel of φ is the set

$$\operatorname{Ker}(\varphi) = \{(g_0, g_2) \text{ st } \iota(g_0) = j(g_2)^{-1}\}. \tag{3.532}$$

Since $\iota(G_0)$ and $j(G_2)$ have no intersections[41] (a part the identity), we have that the kernel reduces to the identity:

$$\operatorname{Ker}(\varphi) = \{e\}. \tag{3.533}$$

The differentials provide the diagram

$$\begin{tikzcd} \mathcal{G}_0 \arrow[r, "(d\iota)_{e_0}"] & \mathcal{G}_1 \arrow[r, shift left, "(d\pi)_{e_1}"] & \mathcal{G}_2 . \arrow[l, shift left, "(dj)_{e_2}"] \end{tikzcd} \tag{3.534}$$

We have $(d\pi)_{e_1} \circ (dj)_{e_2} = \mathrm{id}|_{\mathcal{G}_2}$ and the map

$$(d\varphi)_e \colon \mathcal{G}_1 \to \mathcal{G}_0 \oplus \mathcal{G}_2 \tag{3.535}$$

is an algebra homomorphism (as differential of group homomorphism). It is also an isomorphism by dimension counting. The inverse theorem then shows that φ is a local diffeomorphism: $\varphi(G)$ contains a neighborhood of the identity and then is surjective by proposition 2.1.

We conclude that φ is a Lie group isomorphism.

[41] They are transverse because $j \circ \pi = \mathrm{id}|_{G_2}$.

3.19.3.2 Example: extensions of the Heisenberg algebra

Let $\mathscr{H}(V,\Omega) = V \oplus \mathbb{R}E$ be the Heisenberg algebra. A derivation is a map $D\colon \mathscr{H} \to \mathscr{H}$ such that

$$D[X,Y] = [DX, YT] + [X, DY]. \tag{3.536}$$

Let us look at the derivations under the form

$$D = \begin{pmatrix} X & v \\ \xi & a \end{pmatrix} \tag{3.537}$$

where $a \in \mathbb{R}$, $X \in \operatorname{End}(V)$, $v \in V$ and $\xi \in V^*$. The left hand side of the condition (3.536) reads

$$D[w + zE, w' + z'E] = D\big(\Omega(w,w')E\big) = \Omega(w,w')(v + aE). \tag{3.538}$$

Now, using $Dw = Xw + \xi(w)E$ and $D(zE) = v + aE$, the right hand side is

$$\big(\Omega(Xw, v') + \Omega(zv, v') + \Omega(w, Xw') + \Omega(w, z'v)\big)E. \tag{3.539}$$

Equating (3.538) and (3.539) we find $v = 0$ and

$$\Omega(Xw, w') + \Omega(w, Xw') = a\Omega(w,w'). \tag{3.540}$$

If we write it as matrices, we find

$$X^t\Omega + \Omega X = a\Omega. \tag{3.541}$$

The derivations with $a = 0$ form the algebra

$$\mathtt{Der}(\mathscr{H})_0 = \mathfrak{sp}(\Omega, V) \times V^*. \tag{3.542}$$

If $a \neq 0$, we find the symplectic conform group

$$\operatorname{Conf}(V,\Omega) = \{A\colon V \to V \text{ st } \Omega(Av, Aw) = \lambda\Omega(v,w) \text{ with } \lambda \in \mathbb{R}_0^+\}. \tag{3.543}$$

Taking the derivative of the group condition, we find

$$\frac{d}{dt}\Big[\Omega\big(A(t)v, A(t)w\big)\Big]_{t=0} = \frac{d}{dt}\Big[\lambda(t)\Omega(v,w)\Big]_{t=0}, \tag{3.544}$$

which produces the condition (3.540) with $X = \dot{A}$ and $a = \dot{\lambda}$.

(i) If $X = \mathrm{id}$ and $\xi = 0$, then we must have $a = 2$ and we have the derivation

$$H = \mathrm{id}|_V \oplus 2\,\mathrm{id}|_{\mathbb{R}E}. \tag{3.545}$$

(ii) If $\xi = 0$, $a = 0$ and X if exchange the Lagrangian in the decomposition $V = W \oplus \bar{W}$.

3.19.4 Group algebra

Let $\mathcal{A}$ and $\mathcal{B}$ be abelian algebras and $\rho\colon \mathcal{A} \to \mathtt{Der}\,\mathcal{B}$ be a homomorphism. We want to put a group structure on the set $\mathcal{A} \times \mathcal{B}$ in such a way that the Lie algebra of $\mathcal{A} \times \mathcal{B}$ has Lie bracket given by

$$\big[(A + B), (A' + B')\big] = [A, B'] + [B, A'] = \rho(A)B' - \rho(A')B. \tag{3.546}$$

We claim that the group law should be

$$(a,b)(a',b') = (a + a', e^{\rho(a)}b' + b) \tag{3.547}$$

whose inverse is

$$(a,b)^{-1} = (-a, -e^{-\rho(a)b}) \tag{3.548}$$

Indeed, the general form of the commutator is

$$[X,Y] = \frac{d}{dt}\frac{d}{ds}\Big[\,\mathbf{Ad}(X(t))Y(s)\Big]_{\substack{s=0\\t=0}}$$

with respect to the group law. A path in $\mathcal{A} \times \mathcal{B}$ with tangent vector (a,b) is (at, bt). Then

$$\begin{aligned} \big[(a,b),(a',b')\big] &= \frac{d}{dt}\frac{d}{ds}\Big[(at,bt)(a's,b's)(at,bt)^{-1}\Big]_{\substack{s=0\\t=0}} \\ &= \big(0, -\rho(a)b + \rho(a)b'\big). \end{aligned} \tag{3.549}$$

3.20 Pyatetskii-Shapiro structure theorem

Definition 3.251.
*A **normal** j**-algebra** is a triple* $(\mathfrak{s}, \alpha, j)$ *where*

(i) the Lie algebra $\mathfrak{s}$ *is solvable and such that* $\operatorname{ad}(X)$ *has only real eigenvalues for every* $X \in \mathfrak{s}$*,*

(ii) the map $j\colon \mathfrak{s} \to \mathfrak{s}$ *is an endomorphism of* $\mathfrak{s}$ *such that* $j^2 = -1$ *and*

$$[X, Y] + j[jX, Y] + j[X, jY] - [jX, jY] = 0 \tag{3.550}$$

for every $X, Y \in \mathfrak{s}$*,*

(iii) α *is is a linear form on* $\mathfrak{s}$ *such that*

(a) $\alpha\big([jX, X]\big) > 0$ *if* $X \neq 0$*,*

(b) $\alpha\big([jX, jY]\big) = \alpha\big([X, Y]\big)$*.*

If $\mathfrak{s}'$ is a subalgebra of $\mathfrak{s}$ which is invariant under j, then the triple $(\mathfrak{s}', \alpha|_{\mathfrak{s}'}, j|_{\mathfrak{s}'})$ is a also normal j-algebra and is said to be a **normal** j**-subalgebra** of $\mathfrak{s}$.

A normal j-algebra has a real inner product defined by the formula

$$g(X, Y) = \alpha\big([jX, Y]\big). \tag{3.551}$$

If $\mathfrak{g}$ is an Hermitian Lie algebra[42], we can build a normal j-algebra out of $\mathfrak{g}$ in the following way. First, we choose an Iwasawa decomposition

$$\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{n} \oplus \mathfrak{k}, \tag{3.552}$$

and we pick $\mathfrak{s} = \mathfrak{a} \oplus \mathfrak{n}$. Let $G = ANK$ be the group associated with the Iwasawa decomposition (3.552). The manifold $M = G/K$ is an Hermitian symmetric space, and we have a global diffeomorphism

$$\begin{aligned} R = AN &\to G/K \\ g &\mapsto gK \end{aligned} \tag{3.553}$$

which endows the group R with an exact left invariant symplectic structure and a compatible complex structure, see section **??**. We define α by $\Omega_e = d\alpha$ (Ω is exact) and j is the complex structure evaluated at identity.

A normal j-algebra build from an Hermitian symmetric space of rank 1 (i.e. $\dim \mathfrak{a} = 1$.) is **elementary**. Elementary normal j-algebra are well understood by the following proposition.

Proposition 3.252.
An elementary normal j*-algebra is a split extension*

$$\mathfrak{s}_{el} = \mathfrak{a}_1 \oplus_{\text{ad}} \mathfrak{n}_1 = \mathfrak{a}_1 \oplus_{\text{ad}} \big(V \oplus \mathfrak{z}_1\big) \tag{3.554}$$

where $\mathfrak{n}_1$ *is an Heisenberg algebra* $\mathfrak{n}_1 = V \oplus \mathfrak{z}_1$ *and* $\mathfrak{a}_1$ *is one dimensional. Moreover,* V *is a symplectic vector space and one can choose* $H \in \mathfrak{a}_1$ *and* $E \in \mathfrak{z}_1$ *in such a way that*

$$\begin{aligned} [H, v] &= v, \\ [v, v'] &= \Omega(v, v')E, \\ [H, E] &= 2E. \end{aligned} \tag{3.555}$$

Any normal j-algebra is build from elementary normal j-algebras by mean of the following lemma.

Proposition 3.253.
Let $(\mathfrak{s}, \alpha, j)$*, a normal* j*-algebra and* $\mathfrak{z}_1$*, a one dimensional ideal of* $\mathfrak{s}$*.*

(i) There exists a vector space V *such that*

$$\mathfrak{s}_1 = j\mathfrak{z}_1 + V + \mathfrak{z}_1 \tag{3.556}$$

is an elementary normal j*-algebra, and such that* $\mathfrak{s}$ *is a split extension*

$$\mathfrak{s} = \mathfrak{s}' \oplus_{\text{ad}} \mathfrak{s}_1 \tag{3.557}$$

where $\mathfrak{s}'$ *is, itself, a normal* j*-algebra.*

[42] i.e. the center of its maximal compact is one dimensional.

(ii) If $\mathfrak{s}_1 = \mathfrak{a}_1 \oplus_{\text{ad}} (V \oplus \mathfrak{z}_1)$, then

$$j\mathfrak{z}_1 + \mathfrak{z}_1 = \mathfrak{a}_1 \oplus \mathfrak{z}_1 \tag{3.558}$$

and

$$\begin{aligned}[]
[\mathfrak{s}', \mathfrak{a}_1 \oplus \mathfrak{z}_1] &= 0,\\
[\mathfrak{s}', V] &\subset V.
\end{aligned} \tag{3.559}$$

(iii) Such an ideal $\mathfrak{z}_1$ exists in every normal j-algebra.

Let us see what are the possibilities for j. If $jE = aH + b_i v_i + cE$, then

$$[jE, E] = 2aE. \tag{3.560}$$

We can prove that $a \neq 0$. Indeed, if $a = 0$, then $jE = cE$ and $-E = j^2E = cjE = c^2E$.

Now, we use the following Jacobi identity on $[H, [jE, v]]$ and the commutation relations, we find $b_i = 0$. Now, suppose that $jH = a'H + b'_i + c'E$. In that case,

$$-E = j^2E = j(aH + cE) = aa'H + ab'_i v_i + ac'E + caH + c^2E. \tag{3.561}$$

Since $a \neq 0$, we have $b'_i = 0$. So we have

$$\begin{aligned}
jE &= aH + cE\\
jH &= a'H + c'E.
\end{aligned} \tag{3.562}$$

Expressing that $j^2E = -E$ and $j^2H = -H$, we find the following constrains on the coefficients:

$$\begin{aligned}
aa' + ca &= 0\\
ac' + c^2 &= -1\\
c'^2 + c'a &= -1\\
c'c + c'c &= 0.
\end{aligned} \tag{3.563}$$

We check that $a \neq 0$, $c' \neq 0$ and $a' = -c$. The remaining relation is $c^2 + c'a = -1$. Thus in the basis $\{H, E\}$, the endomorphism j reads

$$j = \begin{pmatrix} -c & a\\ c' & c \end{pmatrix} \tag{3.564}$$

with $\det j = 1$.

Lemma 3.254.
An elementary normal j-algebra has no proper j-ideal.

Proof. Let $\mathfrak{i}$ be a j-ideal of the elementary normal j-algebra $\mathfrak{s}_{el}$. Let $\mathfrak{s}_{el} = \mathfrak{a} \oplus_{\text{ad}} (V \oplus \mathfrak{z})$. We denote by H and E the elements of $\mathfrak{a}$ and $\mathfrak{z}$ (which are one dimensional) who fulfill the standard relations (3.555). If $X = aH + b_i v_i + cE \in \mathfrak{i}$, then $\big[[X, v], v\big] \in \mathfrak{i}$. Using the relations, we conclude that $\mathfrak{z} \subset \mathfrak{i}$. By j-invariance of $\mathfrak{i}$, we have $j\mathfrak{z} \subset \mathfrak{i}$. Now, the fact that $[jE, v] = av$ implies that $\mathfrak{i} = \mathfrak{s}_{el}$. □

The structure of a normal j-algebra $\mathfrak{s}$ is thus as follows. We have the decomposition

$$\mathfrak{s} = \mathfrak{s}' \oplus_{\text{ad}} \Big(\mathfrak{a}_1 \oplus_{\text{ad}} (V_1 \oplus \mathfrak{z}_1)\Big) \tag{3.565}$$

where $\mathfrak{s}'$ is again a normal j-algebra. Furthermore, $\dim \mathfrak{a}_1 = \dim \mathfrak{z}_1 = 1$ and we can choose a basis $H \in \mathfrak{a}_1$, $E \in \mathfrak{z}_1$ such that

$$\begin{aligned}[]
[H, v] &= v\\
[H, E] &= 2E\\
[v, v'] &= \Omega(v, v')E\\
[\mathfrak{s}', V] &\subset V\\
[\mathfrak{s}', \mathfrak{a}_1 \oplus \mathfrak{z}_1] &= 0.
\end{aligned} \tag{3.566}$$

for all $v, v' \in V_1$. The algebra $V_1 \oplus \mathfrak{z}_1$ is an Heisenberg algebra.

The algebra $\mathfrak{s}'$ can be decomposed in the same way again and again up to end up with a sequence of elementary normal j-algebra.

Chapter 4

Fibre bundle

4.1 Vector bundle

Let M be a smooth manifold. A V**-vector bundle** of rank r on M is a smooth manifold F and a smooth projection $p\colon F \to M$ such that

- for any $x \in M$, the fiber $F_x := p^{-1}(x)$ is a vector space of dimension r on the same field that V (let's say $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$).
- for any $x \in M$, there exists an open neighbourhood $\mathcal{U}$ of x and a "chart diffeomorphism" $\phi\colon p^{-1}(\mathcal{U}) \to \mathcal{U} \times V$ such that for any $l \in p^{-1}(y)$,
  - $\phi(l) = (y, \phi_y(l))$
  - $\phi_y\colon E_y \to V$ is a vector space isomorphism.

The pair $(\mathcal{U}, \phi)$ is a *local trivialization*; M is the *base space*; F, the *total space*, p the *projection* and r, the *rank* of the bundle. The denominations of total and base spaces will also be used in the same way for principal bundles.

We will sometimes use charts diffeomorphism $\phi\colon \mathcal{U} \times V \to p^{-1}(\mathcal{U})$ instead of $\phi\colon p^{-1}(\mathcal{U}) \to \mathcal{U} \times V$. Since they are diffeomorphism, this difference don't affect anything.

4.1.1 Transition functions

The trivializations will be denoted by Greek indices: $\mathcal{U}_\alpha$, ϕ_α,... The symbol $\mathcal{U}_{\alpha\beta}$ naturally denotes $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$. If we consider two local trivializations $(\mathcal{U}_\alpha, \phi_\alpha)$ and $(\mathcal{U}_\beta, \phi_\beta)$, we have to look at $\phi_\alpha \circ \phi_\beta^{-1}\colon \mathcal{U}_{\alpha\beta} \times \mathbb{K}^r \to \mathcal{U}_{\alpha\beta} \times \mathbb{K}^r$. We define the **transition functions** $g_{\alpha\beta}\colon \mathcal{U}_{\alpha\beta} \to GL(r, \mathbb{K})$ by

$$\phi_\alpha \circ \phi_\beta^{-1}(x, v) = (x, g_{\alpha\beta}(x)v). \tag{4.1}$$

These functions take their values in $GL(r, \mathbb{K})$ because $\phi_y\colon E_y \to V$ is a vector space isomorphism. Since $(\phi_\alpha \circ \phi_\beta)^{-1} = \phi_\beta \circ \phi_\alpha^{-1}$, it is clear that $g_{\alpha\beta}(x) = g_{\alpha\beta}(x)^{-1}$.

If $x \in \mathcal{U}_{\alpha\beta\gamma} = \mathcal{U}_\alpha \cap \mathcal{U}_\beta \cap \mathcal{U}_\gamma$, we have $\phi_\alpha \circ \phi_\gamma^{-1}(x, v) = (x, g_{\alpha\gamma}(x)v)$, but also $\phi_\alpha \circ \phi_\gamma^{-1} = \phi_\alpha \circ \phi_\beta^{-1}\phi_\beta \circ \phi_\gamma^{-1}$, then

$$(x, g_{\alpha\gamma}(x)v) = (\phi_\alpha \circ \phi_\beta^{-1})(x, g_{\beta\gamma}(x)v) = (x, g_{\alpha\beta}(x)_{\beta\gamma}(x)v). \tag{4.2}$$

Thus $g_{\alpha\gamma}(x) = g_{\alpha\beta}(x)g_{\beta\gamma}(x)$. So, as linear maps, we have

$$g_{\alpha\beta} \circ g_{\alpha\gamma} \circ g_{\gamma\alpha} = \mathbb{1}. \tag{4.3}$$

4.1.2 Inverse construction

Let us consider a manifold M, an open covering $\{\mathcal{U}_\alpha : \alpha \in I\}$ and some functions $g_{\alpha\beta}\colon \mathcal{U}_{\alpha\beta} \to GL(r, \mathbb{K})$ which fulfill relations (4.3). We will build a vector bundle $E \xrightarrow{p} M$ whose transition functions are the $g_{\alpha\beta}$'s. Let $\tilde E$ be the disjoint union

$$\tilde E = \bigsqcup_{\alpha \in I} \mathcal{U}_\alpha \times \mathbb{K}^r,$$

i.e. triples of the form $(x, v, \alpha) \in M \times \mathbb{K}^r \times I$ with the condition that $x \in \mathcal{U}_\alpha$. We define an equivalence relation on $\tilde E$ by $(x, v, \alpha) \sim (y, w, \beta)$ if and only if $x = y$ and $w = g_{\alpha\beta}(x)v$. Next, we define $E = \tilde E / \sim$ and

$\omega\colon \tilde{E} \to E$, the canonical projection. The projection $p\colon E \to M$ is naturally defined by $p([x, v, \alpha]) = x$. The chart diffeomorphism is $\varphi_\alpha\colon \mathcal{U}_\alpha \times \mathbb{K}^r \to p^{-1}(\mathcal{U}_\alpha)$,

$$\varphi_\alpha(x, v) = \omega(x, v, \alpha).$$

Now we have to prove that E endowed with the φ_α's is a vector bundle.

First we prove that φ_α is surjective. For this we remark that a general element in $p^{-1}(\mathcal{U}_\alpha)$ can be written under the form $\omega(x, v, \alpha)$ with $x \in \mathcal{U}_{\alpha\beta}$. But

$$\begin{aligned}\varphi_\alpha(x, g_{\alpha\beta}(x)w) &= \omega(x, g_{\alpha\beta}(x)w, \alpha)\\ &= \omega(x, g_{\alpha\beta}(x)g_{\alpha\beta}(x)w, \beta)\\ &= \omega(x, w\beta),\end{aligned} \tag{4.4}$$

then φ_α is surjective. Now we suppose $\varphi_\alpha(x, v) = \varphi_\alpha(y, w)$. Then $\omega(x, v, \alpha) = \omega(y, w, \alpha)$ and $x = y$, $w = g_{\alpha\alpha}v$ which immediately gives $v = w$. Then φ_α is injective.

Finally, we have

$$(\varphi\alpha \circ \varphi_\beta^{-1})(\omega(x, v, \alpha)) = \varphi_\alpha(x, g_{\alpha\beta}(x)v) = \omega(x, g_{\alpha\beta}(x)v, \alpha), \tag{4.5}$$

which proves that the maps g are the transition functions of the vector bundle E.

4.1.3 Equivalence of vector bundle

Let $E \xrightarrow{p} M$ and $F \xrightarrow{p'} M$ be two vector bundles on M. They are **equivalent** if there exists a smooth diffeomorphism $f\colon E \to F$ such that

- $p' \circ f = p$,
- $f|_{E_x}\colon E_x \to F_x$ is a vector space isomorphism.

Let E and F be two equivalent vector bundles, $\{\mathcal{U}_\alpha \text{ st } \alpha \in I\}$, an open covering which trivialize E and F in the same time and ϕ^E_α, ϕ^F_α the corresponding trivializations. A map $f\colon E \to F$ reads “in the trivialization” as $\phi^F_\alpha \circ f|_{p^{-1}(\mu_\alpha)} \circ \phi^{E}{}^{-1}_\alpha\colon \mathcal{U}_\alpha \times \mathbb{K}^r \to \mathcal{U}_\alpha \times \mathbb{K}^r$ and defines a map $\lambda_\alpha\colon \mathcal{U}_\alpha \to GL(r, \mathbb{K})$ by

$$(\phi^F_\alpha \circ f|_{p^{-1}(\mu_\alpha)} \circ \phi^{E}{}^{-1}_\alpha)(x, v) = (x, \lambda_\alpha(x)v). \tag{4.6}$$

If we denote by g^E the transition functions for E (and g^F for F),

$$\phi^F_\alpha \circ \phi^{F-1}_\beta = (\phi^F_\alpha \circ f \circ \phi^{E-1}_\alpha) \circ (\phi^E_\alpha \circ \phi^{E-1}_\beta) \circ (\phi^E_\beta \circ f^{-1} \circ \phi^{E-1}_\beta),$$

so that

$$g^F_{\alpha\beta}(x) = \lambda_\alpha(x) g^E_{\alpha\beta}(x) \lambda(x)^{-1}. \tag{4.7}$$

Once again we have an inverse construction. We consider a vector bundle E on M with transition functions g^E and some maps $\lambda_\alpha\colon \mathcal{U}_\alpha \to GL(r, \mathbb{K})$; then we define $g^F_{\alpha\beta}(x)$ by equation (4.7).

From subsection 4.1.2, one can construct a vector bundle F on M whose transition functions are these g^F. With the trivializations ϕ^F of F, one can define $f\colon E \to F$ by

$$(\phi^F_\alpha \circ f \circ \phi^{E-1}_\alpha)(x, v) = (x, \lambda_\alpha(x)v).$$

When a basis space B is given, we denote by $\mathrm{Vect}(B)$ the set of isomorphism classes of vector bundles over B. In the complex case, we denote it by $\mathrm{Vect}_\mathbb{C}(B)$.

Proposition 4.1.
Any vector bundle over $\mathbb{R}^n$ is trivial.

Proof. Let $p\colon F \to M$ be a vector bundle on $M = \mathbb{R}^n$ and $\{\mathcal{U}_\alpha\}$ be covering of $\mathbb{R}^n$ by local trivializations. Now consider a partition of unity related to the covering $\mathcal{U}_\alpha$: a set of functions $f_\alpha\colon M \to \mathbb{R}$ such that

- $f_\alpha > 0$,
- $\forall x \in M$, one can find a neighbourhood of x in which only a *finite* number of f_α is non zero,
- $\forall x \in M$, $\sum_\alpha f_\alpha(x) = 1$.
- $f_\alpha = 0$ outside of $\mathcal{U}_\alpha$.

Using that partition of unity, we build the trivialization function $f\colon F \to \mathbb{R}^n \times V$ by $f(l) = (x, \sum_\alpha f_\alpha(x)\phi_{\alpha x}(l))$. □

The following two propositions have some importance in K-theory.

Proposition 4.2.
Let $\pi\colon E \to B$ be a complex vector bundle over a basis compact, Hausdorff, connected basis B. Then there exists a vector bundle E' such that $E \oplus E'$ is trivial.

Proposition 4.3.
Let $f\colon A \to B$ be a map between the topological spaces A and B, and consider a vector bundle $\pi\colon E \to B$. Then there exists one and only one vector bundle $\pi'\colon E' \to A$ and a map $f'\colon E' \to E$ such that $f'|_{E'_x}\colon E'_x \to E_{f(x)}$ is an isomorphism. The vector bundle E' is unique up to isomorphism.

Proofs can be found in [30]. Let us denote by $f^*(E)$ the function given by proposition 4.3. It satisfies the following properties

$$\begin{aligned} (fg)^*(E) &= g^*\big(f^*(E)\big)\\ \mathrm{id}^*(E) &= E\\ f^*(E_1 \oplus E_2) &= f^*(E_1) \oplus f^*(E_2)\\ f^*(E_1 \otimes E_2) &= f^*(E_1) \otimes f^*(E_2). \end{aligned} \tag{4.8}$$

4.1.4 Sections of vector bundle

A **section** of the vector bundle $p\colon E \to M$ is a smooth map $s\colon M \to E$ such that $p \circ s = \mathrm{id}|_M$. The set of all the sections is denoted by $\Gamma^\infty(M)$ or simply $\Gamma(E)$.

If $(\mathcal{U}_\alpha, \phi_\alpha)$ is a local trivialization, one can describe the section s by a function $s_\alpha\colon \mathcal{U}_\alpha \to V$ defined by $\phi_\alpha(s(x)) = (x, s_\alpha(x))$, or equivalently by

$$s(x) = \phi_\alpha^{-1}(x, s_\alpha(x)).$$

As usual when we define such a local quantity, we have to ask ourself how are related s_α and s_β on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$. The best is $s_\alpha = s_\beta$, but most of the time it is not. Here, we compute

$$\phi_\beta \circ \phi_\alpha^{-1} \circ \phi_\alpha(s(s)) = (x, g_{\alpha\beta}(x)s_\alpha(x)),$$

which is obviously also equal to $(x, s_\beta(x))$. Then

$$s_\beta(x) = g_{\alpha\beta}(x)s_\alpha(x) \tag{4.9}$$

without summation.

4.2 Vector valued differential forms

Let E be a vector bundle over M. A **E-valued p-form** is a section

$$e \in \Gamma\big(E \otimes \bigwedge^p T^*M\big).$$

We denote by $\Omega(M, E) = \Gamma\big(E \otimes \bigwedge^p T^*M\big)$ the set of E-valued differential forms. An element of $\Omega^1(M, E) = \Gamma\big(E \otimes \bigwedge T^*M\big)$ always reads $\sum_i s_i \otimes \omega_i$ for some sections s_i and usual differential forms ω_i.

A form of $\Omega^p(M, E)$ can be seen as a fiber morphism $\underbrace{TM \otimes \cdots \otimes TM}_{p \text{ times}} \to E$ by associating

$$s \otimes \omega(X_1, \cdots, X_p) = s(x)\omega(X_1, \cdots, X_p) \in E_x$$

to the element $(s \otimes \omega) \in \Omega^p(M, E)$. There exists a wedge product between vector-valued forms. If $e \in \Omega^p(M, E_1)$ and $f \in \Omega^q(M, E_2)$, then we define $e \wedge f \in \Omega^{p+q}(M, E_1 \otimes E_2)$ by

$$(e \wedge f)(v_1, \cdots, v_{p+q}) = \frac{1}{p!q!} \sum_{\pi \in S_{p+q}} (-1)^\pi e(v_{\pi(1)}, \cdots v_{\pi(p)}) \otimes f(v_{\pi(p+1)}, \cdots, v_{\pi(p+q)}) \in E_1 \otimes E_2. \tag{4.10}$$

where $(-1)^\pi$ stands for the sign of the permutation π. For example when $e, f \in \Omega^1(M, E)$, we have

$$(e \wedge f)(X, Y) = e(X) \otimes f(Y) - e(Y) \otimes f(X) \in E \otimes E.$$

When M is a differentiable manifold, the **fundamental 1-form** is the element $\theta \in \Omega(M, TM)$ such that

$$\iota(X)\theta = X$$

for every $X \in \Gamma(TM)$.

4.3 Lie algebra valued differential forms

An important particular case of vector valued forms is given by Lie algebra valued forms. That case appears for example in the connection theory over principal bundle[1]. If ω and η are elements of $\Omega^1(M, \mathcal{G})$ for some Lie algebra $\mathcal{G}$, we define

$$(\omega \wedge \eta)(X, Y) = \omega(X) \otimes \eta(Y) - \omega(Y) \otimes \eta(X).$$

Combining with the Lie bracket, we define

$$[\omega \wedge \eta](X, Y) := [\omega(X), \eta(Y)] - [\omega(Y), \eta(X)]. \tag{4.11}$$

Using the proposition 7.11, we often implicitly transforms the tensor product into a product (7.41b) and put

$$(\omega \wedge \omega)(X, Y) = [\omega(X), \omega(Y)]. \tag{4.12}$$

Let us point out the fact that that kind of formula only holds for a "wedge square", but not for a general product $\omega \wedge \eta$. Remark that for $\omega \in \Omega^1(M, \mathcal{G})$ and $\beta \in \Omega^2(M, \mathcal{G})$, a simple computation of definition (4.10) yields

$$(\omega \wedge \beta)(X, Y, Z) = \omega(X) \otimes \beta(Y, Z) - \omega(Y) \otimes \beta(X, Z) + \omega(Z) \otimes \beta(X, Y), \tag{4.13}$$

so that, using the same trick as for equation (4.12), we find

$$(\omega \wedge \beta - \beta \wedge \omega)(X, Y, Z) = [\omega(X), \beta(Y, Z)] - [\omega(Y), \beta(X, Z)] + [\omega(Z), \beta(X, Y)].$$

But that expression is exactly what we find by exchanging the tensor product by Lie bracket in expression (4.13). So we define

$$[\omega \wedge \beta] = \omega \wedge \beta - \beta \wedge \omega \tag{4.14}$$

when $\omega \in \Omega^1(M, \mathcal{G})$ and $\beta \in \Omega^2(M, \mathcal{G})$. The reader should remark that this is what one would expect from generalisation of definition (4.11).

4.4 Principal bundle

Let M be a manifold and G, a Lie group whose unit is denoted by e. A G-**principal bundle** on M is a smooth manifold P, a smooth map $\pi\colon P \to M$ and a right action of G on P denoted by $\xi \cdot g$ with $g \in G$ and $\xi \in P$ such that

- $\pi(\xi \cdot g) = \pi(\xi)$,
- $\forall \xi \in \pi^{-1}(x)$, $\pi^{-1}(x) = \{\xi \cdot g \text{ st } g \in G\} \simeq G$,
- $\forall x \in M$, there exists a neighbourhood $\mathcal{U}_\alpha$ of x in M, a diffeomorphism $\phi_\alpha\colon \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times G$ and a diffeomorphism $\phi_{\alpha x}\colon P \to G$ such that
  - $\phi_\alpha(\xi) = (x, \phi_{\alpha x}(\xi))$,
  - $\phi_{\alpha x}(\xi \cdot g) = \phi_{\alpha x}(\xi) \cdot g$.

The group G is often called the **structure group**. We suppose that the action is effective. We will sometimes use the notation $P(G, M)$ to precise that P is a principal bundle over M with structure group G.

The whole construction is given in figure 4.1. All is not yet defined, but in the following, the notations will follow this scheme.

Lemma 4.4.
The map ϕ_α^{-1} fulfills

$$\phi\alpha^{-1}(x, h) \cdot g = \phi_\alpha^{-1}(x, hg).$$

[1] So in Maxwell and other gauge field theories.

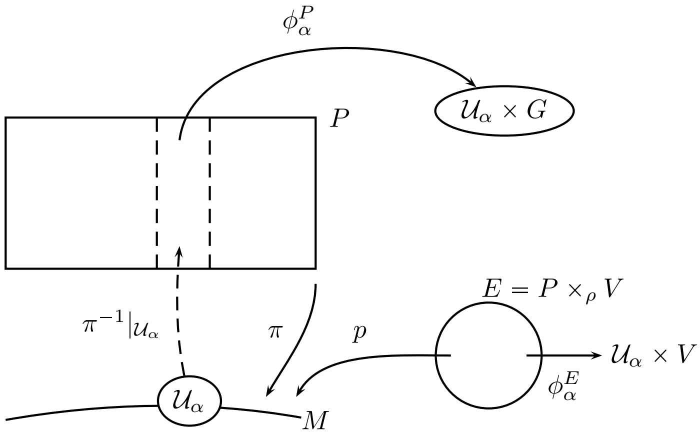


Figure 4.1: Some bundles

Proof. From the definition of a principal bundle, any $\xi \in P$ can be written under the form $\xi = \phi_\alpha^{-1}(x, \phi_{\alpha x}(\xi))$ with ϕ_x satisfying $\phi_x(\xi \cdot h) = \phi_x(\xi)h$ for a certain function $\phi_x : P \to G$. We consider in particular $\xi = \phi_\alpha^{-1}(x,h) \cdot g$. Then $\xi \cdot g^{-1} = \phi_\alpha^{-1}(x,h)$. But $\xi \cdot g^{-1} = \phi_\alpha^{-1}(x, \phi_{\alpha x}(\xi)g^{-1})$, then $h = \phi\alpha_x(\xi)g^{-1}$ and $\phi_{\alpha x}(\xi) = hg$. So we have

$$\xi = \phi_\alpha^{-1}(x,h) \cdot g = \phi_\alpha^{-1}(x, \phi_{\alpha x}(\xi)) = \phi_\alpha^{-1}(x, hg).$$

□

Let

$$R = \{(x,y) \in P \times P \text{ st } x = y \cdot g \text{ for a certain } g \in G\}.$$

Proposition 4.5.
The function $u: R \to G$ defined by the condition

$$p \cdot u(p,q) = q.$$

is differentiable.

Proof. Let $\mathcal{U}$ be an open subset of M and $\sigma: \mathcal{U} \to P$, a section. We consider a differentiable map $\rho: \pi^{-1}(\mathcal{U}) \to G$ such that $\rho(\xi \cdot g) = \rho(\xi) \cdot g$ and $\rho(\sigma(x)) = e$. Such a map is given by

$$\rho(\xi) = \phi_x(\sigma(x))^{-1}\phi_x(\xi)$$

where $x = \pi(\xi)$. We naturally define $R_{\mathcal{U}} = R \cap (\pi^{-1}(\mathcal{U}) \times \pi^{-1}(\mathcal{U}))$ and we pick $(\xi, \eta) \in R_{\mathcal{U}}$. Let $s \in G$ be the one such that $\xi \cdot s = \eta$, so that $\rho(\xi) \cdot s = \rho(\eta)$. Then the restriction of u to $R_{\mathcal{U}}$ is given by $u(\xi,\eta) = \rho(\xi)^{-1}\rho(\eta)$ which makes $u|_{\mathcal{U}}$ differentiable. Since this reasoning can be made on every chart open $\mathcal{U}$, u is differentiable everywhere on P. □

The following is a corollary of Leibnitz rule.

Corollary 4.6.
If P is a G-principal bundle and v, a are curve in P and G respectively, we can consider the curve $u(t) = v(t)a(t)$. We have:

$$\frac{d}{dt}u(t)\Big|_{t=0} = \frac{d}{dt}v(t)a(0)\Big|_{t=0} + \frac{d}{dt}v(0)a(t)\Big|_{t=0}.$$

The proof is direct. This result is often written as

$$\dot{u}_t = \dot{v}_t a_t + v_t \dot{a}_t. \tag{4.15}$$

A main application is

$$\frac{d}{dt}\Big[r \cdot h(t)\Big]_{t=0} = \frac{d}{dt}\Big[r \cdot e^{th'(0)}\Big]_{t=0}. \tag{4.16}$$

4.4.1 Transition functions

Let $(\mathcal{U}_\alpha, \phi_\alpha)$ be a local trivialization of P. This induces transition functions $g_{\alpha\beta}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ defined by

$$\begin{aligned} \phi_\alpha \circ \phi_\beta^{-1}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \times G &\to \mathcal{U}_\alpha \cap \mathcal{U}_\beta \times G \\ (x,a) &\mapsto (x, g_{\alpha\beta}(x)a). \end{aligned} \tag{4.17}$$

Clearly, $g_{\alpha\alpha} = e$ and $g_{\alpha\beta} g_{\alpha\beta} = e$ on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$. Then the triviality

$$\phi_\alpha \circ \phi_\beta^{-1} \circ \phi_\beta \circ \phi_\gamma^{-1} \circ \phi_\gamma \circ \phi_\alpha^{-1} = \mathrm{id}$$

implies the compatibility conditions

$$g_{\alpha\beta} g_{\beta\gamma} g_{\gamma\alpha} = e \tag{4.18}$$

on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \cap \mathcal{U}_\gamma$.

There is an inverse construction. Let $\{\mathcal{U}_\alpha \text{ st } \alpha \in I\}$ be an open covering of M and $g_{\alpha\beta}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ a family of functions such that $g_{\alpha\alpha} = e$, $g_{\alpha\beta} g_{\alpha\beta} = e$ on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ and $g_{\alpha\beta} g_{\beta\gamma} g_{\gamma\alpha} = e$ on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta \cap \mathcal{U}_\gamma$. Then the following construction gives a G-principal bundle whose transition functions are the $g_{\alpha\beta}$'s.

- $\tilde P = \bigsqcup_{\alpha\in I} \mathcal{U}_\alpha \times G$ (disjoint union),
- if $(x,a) \in \mathcal{U}_\alpha \times G$ and $(y,b) \in \mathcal{U}_\beta \times G$, then $(x,a) \sim (y,b)$ if and only if $x = y$ and $b = g_{\alpha\beta}(x)a$,
- $\pi\colon \tilde P \to M$ is defined by $\pi[(x,a)] = x$ where $[(x,a)]$ is the class of (x,a) for $\sim$,
- the action is defined by $[(x,a)] \cdot g = [(x, ag)]$.

Theorem 4.7.
Let G be a Lie group; M, a differentiable manifold; $\{\mathcal{U}_\alpha\}_{\alpha\in I}$, an open covering of M and some functions $\varphi_{\alpha\beta}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$ such that $\varphi_{\alpha\beta}(x) = \varphi_{\alpha\gamma}(x)\varphi_{\gamma\beta}(x)$. Then there exists a principal bundle P whose transition functions are the φ_α's for the covering $\{\mathcal{U}_\alpha\}_{\alpha\in I}$.

Proof. We consider the topological space

$$E = \bigcup_{\alpha\in I} (G \times \mathcal{U}_\alpha \times I) \tag{4.19}$$

where we put the discrete topology on I. Each $G \times \mathcal{U}_\alpha \times \{\alpha\}$ is a manifold. Thus E has a structure of differentiable manifold induced from the one of $G \times M$. We consider on E the equivalence relation given by the following subset of $E \times E$:

$$R = \big\{ \big((g,x,\alpha),(h,y,\beta)\big) \in E \times E \text{ st } y = x \text{ and } h = \varphi_{\alpha\beta}(x) g \big\}.$$

We will show that $P = E/R$ has a structure of principal bundle. We begin by defining an action of G on P by

$$[(g,x,\alpha)\cdot h] = [(gh,x,\alpha)].$$

In order to see that this definition is correct, let us consider $[g',x,\beta] = [g,x,\alpha]$. From the definition of the equivalence class, $g' = \varphi_{\alpha\beta}(x)g$. Then $[(g',x,\beta)]\cdot h = [(\varphi_{\alpha\beta}(g)gh, x, \beta)]$, and the form of R shows that this is well $[(gh,x,\alpha)]$. Since the map $(g,h) \to gh$ is differentiable on G, the so defined action is a differentiable action of G on P and G is a transformation group on P[2].

If $[(g,x,\alpha)] = [(gh,x,\alpha)]$, then $gh = \varphi_{\alpha\alpha} g = g$ and $h = e$. So the action is effective.

Now we consider the quotient P/G. A typical element is

$$\overline{(s,x,i)} = \{[s,x,i]\cdot g \text{ st } g \in G\}.$$

The projection $\pi\colon P \to M$, $[(s,x,\alpha)] \to x$ is well defined and we can consider $\varphi\colon P/G \to M$, $\varphi\overline{(s,x,\alpha)} = x$. It provides a bijection between P/G and M. So we can identify P/G and M. Now we are going to show that P endowed with the projection $\pi\colon P \to X$ is a principal bundle.

We consider the map

$$h_\alpha\colon G \times \mathcal{U}_\alpha \to P \ (g,x) \quad \mapsto \omega(g,x,\alpha) \tag{4.20}$$

where $\omega\colon E \to P = E/R$ is the canonical projection. Since

$$(\pi \circ h_\alpha)(g,x) = (\pi\circ\omega)(g,x,\alpha) = \pi[(g,x,\alpha)] = x,$$

[2] Faut voir comment ça correspond à la définition de l'autre texte.

the map h_α actually is $h_\alpha\colon G\times\mathcal{U}_\alpha\to\pi^{-1}(\mathcal{U}_\alpha)$. In order to see that h_α is surjective on $\pi^{-1}(\mathcal{U}_\alpha)$, let us take a general element of $\pi^{-1}(\mathcal{U}_\alpha)$ under the form $\omega(g,x,\beta)$ with $x\in\mathcal{U}_\alpha\cap\mathcal{U}_\beta$. Then $(g,x,\beta)\in[(\varphi_{\alpha\beta}(x)g,x,\alpha)]$ and therefore $\omega(g,x,\beta)=h_\alpha(\varphi_{\alpha\beta}(x)g,x)$. For the injectivity, remark that $\omega(g,x,\beta)=\omega(h,y,\alpha)$ implies $x=y$ and $h=\varphi_{\beta\beta}(x)g=g$. In particular, $h_\alpha(g,x)=h_\alpha(h,y)$ implies $x=y$ and $g=h$.

Now we will prove that the inverse of h_α is continuous. For this we consider an open set $\Omega\subset G\times\mathcal{U}_\alpha$ and we have to show that $h_\alpha(\Omega)$ is open in $\pi^{-1}(\mathcal{U}_\alpha)$.

We recall the **quotient topology**: if A is a topological space with an equivalence relation $\sim$ and the canonical projection $\varphi\colon A\to A/\sim$, then $V\subset A/\sim$ is open if and only if $\varphi^{-1}(V)\subset A$ is open. So in our case, we have to check the openness of $V=\omega^{-1}(h_\alpha(\Omega))$ in E. We consider the open covering

$$\{G\times\mathcal{U}_\alpha\times\{\alpha\}\}_{\alpha\in I}$$

of E and we will show that the intersection of V with any of these open set is open. We have to show that $\omega^{-1}\big(h_\alpha(\Omega)\cap(G\times\mathcal{U}_\alpha\times\{\beta\})\big)$ is open for any $\beta\in I$. For this, we define a map $\alpha\colon G\times(\mathcal{U}_\alpha\cap\mathcal{U}_\beta)\times\{\beta\}\to G\times\mathcal{U}_\alpha$ by

$$\alpha_\beta(g,x,\beta)=(\varphi_{\alpha\beta}(x)g,x) \tag{4.21}$$

which is continuous. The set $(h_\alpha\circ\alpha_\beta)^{-1}(h_\alpha(\Omega))=\alpha_\beta^{-1}(\Omega)$ is open because $h_\alpha\circ\alpha_\beta$ is the restriction of ω to $G\times(\mathcal{U}_\alpha\cap\mathcal{U}_\beta)\times\{\beta\}$. Then h_α is an homeomorphism from $G\times\mathcal{U}_\alpha$ tp $\pi^{-1}(\mathcal{U}_\alpha)$. Since it is build from differentiable functions, it is moreover a diffeomorphism.

So we have a chart system $\{(h_\alpha,\mathcal{U}_\alpha)\}_{\alpha\in I}$ where h_α fulfils the "good" properties with respect to π. It remains to be proved that the $\varphi_{\alpha\beta}$'s are the transition functions and that $\pi^{-1}(\pi(\xi))=\xi\cdot G$ for every $\xi\in P$. We begin by the latter. For $\xi=[(g,x,\alpha)]$, $\pi(\xi)=x$ and we have to study the set

$$\pi^{-1}(x)=\{[(h,x,\beta)]\text{ st }h\in G,\beta\in I\}.$$

Clearly, $[(h,x,\beta)]\cdot G\subset\pi^{-1}(x)$. The fact that there is nothing else than $[(h,x,\beta)]\cdot G$ in $\pi^{-1}(x)$ is seen by

$$[h,x,\beta]=[\varphi_{\alpha\beta}(x)g,x,\alpha]\in[(h,x,\alpha)]\cdot G.$$

In order to check the change of charts, let us consider $g'=h_{\beta,x}^{-1}\circ h_{\alpha,x}(g)$ where

$$h_{\alpha,x}(g)=h_\alpha(g,x)=\omega(g,x,\alpha). \tag{4.22}$$

The fact that $h_\beta(g',x)=g_\alpha(g,x)$ concludes the proof. To see this fact, remark that $h_{\beta,x}(h_{\beta,x}^{-1}\circ h_{\alpha,x}(g))=h_{\alpha,x}(g)$, so that $h_\alpha(g',x)=h_\alpha(g,x)$ implies $\omega(g',x,\beta)=\omega(g,x,\alpha)$ which proves that $g'=\varphi_{\alpha\beta}(g)$. □

The **trivial bundle** is simply $P=M\times G$ and $\pi(x,g)=x$ with the action $(x,a)\cdot g=(x,ag)$.

4.4.2 Morphisms and such...

An **homomorphism** between $P(G,M)$ and $P'(G',M')$ is a differentiable map $h\colon P\to P'$ such that $\forall\xi\in P,g\in G$,

$$h(\xi\cdot g)=h(\xi)\cdot h_G(g) \tag{4.23}$$

where $h_G\colon G\to G'$ is a Lie group homomorphism. From the definition, h maps a fiber to only one fiber, but it is not specially surjective on any fiber. So h induces a homomorphism $h_M\colon M\to M'$ such that $\pi'\circ h=h_M\circ\pi$.

An **isomorphism** is a homomorphism $g\colon P(G,M)\to P'(G',M')$ such that

- h_P is a diffeomorphism $P\to P'$,
- h_G is a Lie group homomorphism $G\to G'$, and
- h_M is a diffeomorphism $M\to M'$.

A principal bundle is **trivial** if one can find an isomorphism $h\colon G\times M\to P$ such that $\pi\circ h=\mathrm{id}\circ\mathtt{pr}_2$, i.e. the following diagram commutes:

$$\begin{tikzcd} G\times M \arrow[r, "h"] \arrow[d, "\mathtt{pr}_2"] & P \arrow[d, "\pi"] \\ M \arrow[r, "\mathrm{id}"'] & M \end{tikzcd} \tag{4.24}$$

We say that P is **locally trivial** if for every $x\in M$, there exists an open neighbourhood $\mathcal{U}$ in M such that $\pi^{-1}(\mathcal{U})$ endowed with the induced structure of principal bundle is trivial.

4.4.3 Frame bundle: first

In the ideas, the building of a vector bundle is just to put a vector space on each point of the base manifold. A principal bundle is to put something on which a group acts on each point. If you have a vector bundle on a manifold, you can consider, on each point $x \in M$, the set of all the basis of the fiber E_x over x. The group $GL(r, \mathbb{K})$ naturally acts on this set which becomes a candidate to be a $GL(r, \mathbb{K})$-principal bundle.

More formally, we consider a vector bundle $F \xrightarrow{p} M$, and for each x, the set of the basis of the vector space $F_x = p^{-1}(x)$. We define

$$P = \bigcup_{x \in M} (\text{basis of } F_x).$$

We naturally consider the projection $\pi \colon P \to M$, $\pi(b_x) = x$ if b_x is a basis of F_x.

Let $\phi^F_\alpha \colon p^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times \mathbb{K}^r$ be a local trivialization of F, and $\{\overline{e}_1, \ldots, \overline{e}_r\}$, the canonical basis of $\mathbb{K}^r$. We naturally define

$$\overline{S}_{\alpha i}(x) = \phi^{F\,-1}_\alpha(x, \overline{e}_i).$$

The set $\{\overline{S}_{\alpha 1}(x), \ldots, \overline{S}_{\alpha r}(x)\}$ is a "reference" basis of F_x with respect to the trivialization ϕ_α. If we choose another basis $\{\overline{v}_1, \ldots \overline{v}_r\}$ of F_x, we can find a matrix $A \in GL(r, \mathbb{K})$ such that $\overline{v}_k = A^l_k \overline{S}_{\alpha l}(x)$. This gives a bijection

$$\begin{aligned} \phi^P_\alpha \colon \pi^{-1}(\mathcal{U}_\alpha) &\to \mathcal{U}_\alpha \times GL(r, \mathbb{K}) \\ (\overline{v}_1, \ldots, \overline{v}_r) &\mapsto (x, A). \end{aligned} \tag{4.25}$$

One can give to P a $GL(r, \mathbb{K})$-principal bundle structure such that the ϕ^P_α are diffeomorphism.

Let $(\mathcal{U}_\alpha, \phi^F_\alpha)$ be a local trivialization of F and $g^F_{\alpha\beta} \colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to GL(r, \mathbb{K})$. In this case, $(\mathcal{U}_\alpha, \phi^P_\alpha)$ is a trivialization of P whose transition function is $g^P_{\alpha\beta} = g^F_{\alpha\beta}$. Indeed

$$\phi^P_\alpha \circ \phi^{P\,-1}_\beta(x, A) = \phi^P_\alpha(\{\overline{v}_1, \ldots, \overline{v}_r\})$$

where $\overline{v}_s = (\phi^F_\beta)^{-1}(x, A^l_s \overline{e}_l)$. In order to see it, recall that $\overline{v}_s = A^l_s \overline{S}_{\alpha l}(x)$ and that $\phi^{F\,-1}_\alpha(x, \overline{e}_s) = \overline{S}_{\alpha s}(x)$. Then

$$\overline{v}_s = (\phi^F_\beta)^{-1}(x, A^l_s \overline{e}_l) = A^l_s \overline{S}_{\alpha s}(x).$$

On the other hand, from the definition of ϕ^P_β, the basis $(\phi^P_\beta)^{-1}(x, A)$ is the one obtained by applying A on S. With all this,

$$\begin{aligned} \phi^P_\alpha \circ (\phi^P_\beta)^{-1}(x, A) &= \phi^P_\alpha\{(\phi^F_\beta)^{-1}(x, A^l_s \overline{e}_l)\}_{s=1,\ldots r} \\ &= \phi^P_\alpha\{(\phi^F_\alpha)^{-1} \circ (\phi^E_\alpha \circ \phi^{F\,-1}_\beta)(x, A^l_s \overline{e}_l)\}_{s=1,\ldots r} \\ &= \phi^P_\alpha\{(\phi^E_\alpha)^{-1}(x, g^F_{\alpha\beta}(x)^s_i A^l_s \overline{e}_l)\}_{i=1,\ldots r} \\ &= (x, g^F_{\alpha\beta}(x) A). \end{aligned} \tag{4.26}$$

The last product $g^F_{\alpha\beta}(x)A$ is a matricial product.

4.4.4 Frame bundle: second

4.4.4.1 Basis

If M is a m-dimensional manifold, a **frame** of T_xM is an isomorphism $b \colon \mathbb{R}^m \to T_xM$. In our purpose, we will always deal with (pseudo)Riemannian manifold. So, the tangents spaces T_xM comes with a metric, and we ask a frame to be isometric. In other words, we ask b to be an isometry from $(\mathbb{R}^m, \cdot)$ to (T_xM, g_x), where the dot denotes the (pseudo)euclidian product on $\mathbb{R}^m$. Such a frame is given by a base point x of M and a matrix S in $\mathrm{SO}(g_x)$:

$$b(v) = (Sv)^i (\partial_i)_x, \tag{4.27}$$

if the vector v is written as $v = v^i \overline{1}_i$ in the canonical orthogonal frame $\{\overline{1}_i\}$ of $\mathbb{R}^m$ and $\mathrm{SO}(g_x)$ is the set of the $m \times m$ matrix A such that $A^t g_x A = g_x$.

This frame intuitively corresponds to the basis of T_xM (see as a "true" vector space) that we would have written by $\{Se_i\}_x$ if $e_i = \frac{\partial}{\partial x^i}$. In order to follow this idea, we will effectively denote by $\{Se_i\}_x$ the map $b \colon \mathbb{R}^m \to T_xM$ given by (4.27).

We will often write the frame b as $\{be_i\}_x$, making no differences in notation between the b of $\mathrm{SO}(M)$ and the b of $\mathrm{SO}(g_x)$ which implement it.

Remark 4.8.
One has to distinguish a frame *and a* basis*: a basis is only a free and generator set while a frame can be interpreted as an ordered basis.*

4.4.4.2 Construction

We just saw how to build a frame bundle over a manifold. One can get another expression of the frame bundle when we express a basis of T_xM by means of an isomorphism between $\mathbb{R}^n$ and T_xM. If M is a n-dimensional manifold, a **frame** at x is an ordered basis

$$b = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$$

of T_xM. It is clear that any frame defines an isomorphism (linear bijective map)

$$\begin{aligned}\tilde{b}\colon \mathbb{R}^n &\to T_xM \\ e_i &\mapsto \mathbf{e_i}\end{aligned} \tag{4.28}$$

where $\{e_i\}$ is the canonical basis of $\mathbb{R}^n$. It is also clear that any isomorphism gives rise to a frame. Then we see a frame of M at x as an isomorphism $\tilde{b}\colon \mathbb{R}^n \to T_xM$. Let $B(M)_x$ be the of all the frames of M at x; we define

$$B(M) = \bigcup_{x\in M} B(M)_x.$$

For all $b \in B(M)_x$, we define $p_B(b) = x$ and the action $B(M) \times GL(n,\mathbb{R}) \to B(M)$ by $b \cdot g = (\mathbf{b}'_1, \ldots, \mathbf{b}'_n)$ where

$$\mathbf{b}'_j = \mathbf{b}_i {g_i}^j. \tag{4.29}$$

It is easy to see that $\widetilde{b \cdot g} = \tilde{b} \circ g\colon \mathbb{R}^n \to T_xM$. So we can give to

$$\begin{tikzcd} GL(n,\mathbb{R}) \arrow[r, squiggly] & B(M) \arrow[d, "p_B"] \\ & M \end{tikzcd} \tag{4.30}$$

a structure of principal bundle[3]. If $(\mathcal{U}_\alpha, \varphi_\alpha)$ is a local coordinate chart on M, we define

$$\tilde{\varphi}\colon p_B^{-1}(\mathcal{U}_\alpha) \to \varphi_\alpha(\mathcal{U}_\alpha) \times GL(n,\mathbb{R})\ b \quad\mapsto (\varphi_\alpha(x), A(b)) \tag{4.31}$$

where $A(b) \in GL(n,\mathbb{R})$ is defined by the condition $\mathbf{b}_j = {A_j}^i \partial_i|_x$. The matrix $A(b)$ is the one which transforms the canonical basis (in the trivialization φ_α) into $b \in B(M)_x$. That's for the principal bundle structure.

The manifold structure of $B(M)$ is given by $\Phi_\alpha\colon p_B^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times GL(\mathbb{R})$,

$$\begin{aligned}\Phi(b) &= (\varphi_\alpha^{-1} \times \mathrm{id}\,|_{GL(n,\mathbb{R})}) \circ \tilde{\varphi}(b) \\ &= (x, A(b)) \\ &= (p_B(b), A(b)).\end{aligned} \tag{4.32}$$

It fulfils $A(b \cdot g) = A(b) \cdot g$. A section $s\colon \mathcal{U}_\alpha \to B(M)$ is sometimes called a **moving frame** over $\mathcal{U}_\alpha$.

Frame bundle over $\mathbb{R}^2$ is given as example in page 203

4.4.5 Sections of principal bundle

A **section** of a G-principal bundle is a smooth map $s\colon M \to P$ such that $s(x) \in \pi^{-1}(x)$ for any $x \in M$. A trivialization ϕ_α^P P on $\mathcal{U}_\alpha$ defines a section of P over $\mathcal{U}_\alpha$ by

$$\sigma_\alpha(x) = (\phi_\alpha^P)^{-1}(x, e)$$

where e is the neutral of the group. In the inverse sense, we have the following:

Proposition 4.9.
If $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$ is local section of P over $\mathcal{U}_\alpha \subset M$, then the definition $\phi_\alpha^P(\xi) = (x, a)$ if $\xi = \sigma_\alpha(x) \cdot a$ is a local trivialization.

Proof. The function ϕ_α^P is well defined because $\xi \in \pi^{-1}(\mathcal{U}_\alpha)$ implies the existence of a $x \in \mathcal{U}_\alpha$ such that $\xi \in \pi^{-1}(x) = \{\xi \cdot g\} \simeq G$. For this x, there exists a $g \in G$ such that $\xi = \sigma_\alpha(x) \cdot g$.

Now we prove that the couple (x, a) is unique in the sense that $s_\alpha(x) \cdot a = \sigma_\alpha(y) \cdot b$ implies $(x, a) = (y, b)$. The left hand side belongs to $\pi^{-1}(x)$ while the right one belongs to $\pi^{-1}(y)$. Then $x = y$. The condition $\pi^{-1}(x) \simeq G$ imposes the unicity of the g making $\xi = \eta \cdot g$ for each couple, $\xi, \eta \in \pi^{-1}(x)$. ∎

[3]Much more details and proofs are given in [31].

If σ and σ' are two sections of the same principal bundle P, then there exists a differentiable map $f\colon M \to G$ such that $\sigma'(x) = \sigma(x) \cdot f(x)$. So all the sections can be deduced from only one and multiplication by such a f.

Theorem 4.10.
If $\pi\colon P(G, M) \to M$ is a principal bundle, then the four following propositions are equivalent:

(i) P is trivial,

(ii) P has a global section,

(iii) there exists a differentiable map $\gamma\colon P \to G$ such that $\gamma(\xi \cdot g) = g^{-1}\gamma(\xi)$ for all $\xi \in P$ and $g \in G$,

(iv) there exists a differentiable map $\rho\colon P \to G$ such that $\rho(\xi \cdot g) = \rho(\xi)g$.

Proof. *(i)⇒ (ii).* The diagram (4.24) commutes and

$$\tau\colon M \to G \times M\; x \quad \mapsto (e, x) \tag{4.33}$$

is a local section of $G \times M$. From it we build the following global section of P:

$$\sigma\colon M \to P\; x \quad \mapsto h(e, x). \tag{4.34}$$

This is injective because $\pi \circ h = \mathtt{pr}_2$ and differentiable because this is a composition of $x \to (e, x)$ and $(g, x) \to h(g, x)$.
(ii)⇒ (i). The principal bundle P admits a global section $\sigma\colon M \to P$. From it, we can build the differentiable map

$$h\colon G \times M \to P\; (g, x) \quad \mapsto \sigma(x) \cdot g \tag{4.35}$$

which satisfies $h(gh, x) = h(g, x) \cdot h$ and $\pi \circ (g, x) = x$. First we show that h is a fiber homomorphism and an isomorphism between P and $G \times M$ so that P is trivial. For this remark that

$$g(gh, x) = g(g, x) \cdot h = \sigma(x) \cdot gh,$$

hence equation (4.23) reduces to $h((g, x)\cdot h) = h(g, x)\cdot h_G(h)$ which is true with $h_G = \mathrm{id}$. Moreover $h\colon G\times M \to P$ is bijective because $\sigma(\pi(\xi))$ belongs to the fiber of $\xi \in P$, therefore there is one and only one $\gamma(\xi) = u(\xi, \sigma(\pi(\xi)))$ such that $\xi \cdot \gamma(\xi) = (\sigma \circ \pi)\xi$. The inverse map is

$$\theta\colon P \to G \times M\; \xi \quad \mapsto (\gamma(\xi), \pi(\xi)) \tag{4.36}$$

which is differentiable because γ and π are. So far we see that h and h^{-1} are differentiable. Then h is an isomorphism between P and $G \times M$.
(ii)⇒ (iii). Let σ be the global section and define

$$\gamma\colon P \to G\; \xi \quad \mapsto u(\xi, (\sigma \circ \pi)\xi) \tag{4.37}$$

where $u\colon R \to G$ is the map defined by the condition $\xi \cdot (\xi, \eta) = \eta$. The map γ is differentiable and we have to prove that $\gamma(\xi \cdot g) = g^{-1}\gamma(\xi)$. Since $\xi \cdot \gamma(\xi) = \sigma \circ \pi(\xi)$,

$$\gamma(\xi \cdot g) = u(\xi \cdot g, (\sigma \circ \pi)(\xi \cdot g)) = u(\xi \cdot g, (\sigma \circ \pi)(\xi)).$$

But $(\xi \cdot g)(g^{-1}\gamma(\xi)) = \xi \cdot \gamma(\xi) = x$. So $\gamma(\xi \cdot g) = u(\xi \cdot g, x)$. Thus $\gamma(\xi \cdot g) = g^{-1}\gamma(\xi)$.
(iii)⇒ (ii). The given map γ fulfils $\xi \cdot g\gamma(\xi \cdot g) = \xi \cdot (\xi)$, so

$$\varphi\colon P \to P\; \xi \quad \mapsto \xi \cdot (\xi) \tag{4.38}$$

is just function of the class of ξ, thus we have a section $\sigma'\colon P/G \to P$, but we know that P/G and M are isomorphic.
(iii)⇒ (iv). Let us define $\rho\colon P \to G$ by $\rho = J \circ \gamma$ with $J(g) = g^{-1}$, thus $\rho(\xi) = \gamma(\xi)^{-1}$ and

$$\rho(\xi \cdot g) = \gamma(\xi \cdot g)^{-1} = (g^{-1}\gamma(\xi))^{-1} = \gamma(\xi)^{-1}g = \rho(\xi)g.$$

(iv)⇒ (iii). The proof is just the same with $\rho = J \circ \rho$. □

Definition 4.11.
*A section $\psi \in \Gamma(P, TP)$ is G-**equivariant** when*

$$d\tau_g\psi(\xi) = \psi(\xi \cdot g).$$

Be careful: this *does not* define equivariant sections of the principal bundle.

4.4.6 Equivalence of principal bundle

Two principal bundles $\pi\colon P \to M$ and $\pi'\colon P' \to M$ are **equivalent** if there exists a diffeomorphism $\varphi\colon P \to P'$ such that

- $\pi' \circ \varphi = \pi$
- $\varphi(\xi \cdot g) = \varphi(\xi) \cdot g$.

If $\{\mathcal{U}_\alpha\}_{\alpha\in I}$ is an open covering of M on which we have trivializations ϕ_α of P and ψ_α of P', the diffeomorphism φ induces some functions $\lambda\colon \mathcal{U}_\alpha \to G$ by setting

$$(\phi_\alpha \circ \varphi^{-1} \circ \psi_\alpha^{-1})(x, a) = (x, \lambda_\alpha(x)a).$$

This definition works because from the definitions of principal bundle and equivalence, one sees that $(\phi_\alpha \circ \varphi^{-1} \circ \psi_\alpha^{-1})(x, \cdot) = (x, \cdot)$.

4.4.6.1 Transition functions

We have some transition functions for P and P' given by equations

$$\begin{aligned}
(\phi_\alpha \circ \phi_\beta^{-1})(x, g) &= (x, g_{\alpha\beta}(x)g)\\
(\psi_\alpha \circ \psi_\beta^{-1})(x, g) &= (x, g'_{\alpha\beta}(x)g).
\end{aligned}$$

Now, we want to know what is $g'_{\alpha\beta}$ in function of $g_{\alpha\beta}$. First remark that $(\psi_\alpha \circ \varphi \circ \phi_\alpha^{-1})(x, a) = (x, \lambda_\alpha(x)^{-1})a$, and next, compute

$$\begin{aligned}
(x, g_{\alpha\beta}(x)a)a &= (\psi_\alpha \circ \varphi \circ \phi_\beta^{-1} \circ \phi_\beta \circ \varphi^{-1} \circ \psi_\beta^{-1})(x, a)\\
&= (\psi_\alpha \circ \varphi \circ \phi_\beta^{-1})(x, \lambda_\beta(x)a)\\
&= (\psi_\alpha \circ \varphi \circ \phi_\alpha^{-1} \circ \phi_\alpha \circ \phi_\beta^{-1})(x, \lambda_\beta(x)a)\\
&= (x, \lambda_\alpha(x)^{-1} g_{\alpha\beta}(x)\lambda_\beta(x)a).
\end{aligned} \tag{4.39}$$

Then

$$g_{\alpha\beta}(x) = \lambda_\alpha(x)^{-1} g_{\alpha\beta}(x)\lambda_\beta. \tag{4.40}$$

One can show that if two principal bundle have transition functions whose fulfill this condition, they are equivalent. A G-principal bundle is **trivial** if it is equivalent to the one given by $\pi_1\colon M \times G \to M$.

4.4.7 Reduction of the structural group

We say that a principal bundle $P(G, M)$ is **reducible** when there exists a principal bundle $P'(H, M)$ such that

- H is a subgroup of G,
- there exists an homeomorphism $h\colon P' \to P$ such that $h_G\colon H \to G$ is an injective homomorphism.

In this case we say that G is reducible to H and that P' is a reduced principal bundle.

Theorem 4.12.
If P is a principal bundle over M, the structural group G is reducible to the Lie subgroup H if and only if there exists an open covering $\{\mathcal{U}_i\}_{i\in I}$ of M and transition functions φ_{ij} taking their values in H.

Proof. No proof. ☐

The following comes from [32]. Let us consider the principal bundle

$$\begin{tikzcd}
G \arrow[r, squiggly] & P \arrow[d, "\pi_P"] \\
& M
\end{tikzcd} \tag{4.41}$$

and H, a closed subgroup of G. We denote by $j\colon H \to G$ the inclusion map. The principal bundle

$$\begin{tikzcd}
H \arrow[r, squiggly] & Q \arrow[d, "\pi_Q"] \\
& M
\end{tikzcd} \tag{4.42}$$

is a **reduction** of P to the group H if there exists a map $u\colon Q\to P$ such that $\pi_P\circ u=\pi_Q$ and $u(\xi\cdot h)=u(\xi)\cdot j(h)$. In this case, u is an embedding[4] of Q in P and the image is a closed submanifold of P.

Let M be a n-dimensional manifold and $B(M)$ be its frame bundle. This is a $\mathrm{GL}(n,\mathbb{R})$-principal bundle. If G is a closed subgroup[5] of $\mathrm{GL}(n,\mathbb{R})$, a G**-structure** is a reduction of $B(M)$ to G.

4.4.8 Density

A **density** on a d-dimensional manifold M is a section of the principal bundle whose fiber P_x over $x\in M$ is the space of homogeneous non vanishing maps

$$\rho\colon \bigwedge^d T_xM\to\mathbb{R}^*_+ \tag{4.43}$$

such that $\rho(\lambda v)=|\lambda|\rho(v)$ for every $\lambda\in\mathbb{R}$ and $v\in\bigwedge^d T_xM$.

4.5 Associated bundle

Let $\pi\colon P\to M$ be a G-principal bundle and $\rho\colon G\to GL(V)$, a representation of G on a vector space V (on $\mathbb{K}=\mathbb{R}$ or $\mathbb{C}$) of dimension r.

The associated bundle $E=P\times_\rho V\xrightarrow{p}M$ is defined as following. On $P\times V$, we consider the equivalence relation

$$(\xi,v)\sim(\xi\cdot g,\rho(g^{-1})v)$$

for $g\in G$, $\xi\in P$ and $v\in V$. Then we define

- $E=P\times_\rho V:=P\times V/\sim$,
- $p[(\xi,v)]=\pi(\xi)$

where $[(\xi,v)]$ is the class of (ξ,v) in $P\times V$.

If $\phi^P_\alpha(\xi)=(\pi(\xi),a(\xi))$ is a trivialization of P on $\mathcal{U}_\alpha$, then

$$\phi^E[(\xi,v)]=(\pi(\xi),\rho(a)v) \tag{4.44}$$

is a trivialization of E.

In order to see that it is a good definition, let us consider $(\eta,w)\sim(\xi,v)$. It immediately gives the existence of a $g\in G$ such that $\eta=\xi\cdot g$ and $w=\rho(g^{-1})v$. Then $\phi^E[(\xi\cdot g,\rho(g^{-1})v)]=(\pi(\xi\cdot g),\rho(b)\rho(g^{-1})v)$. From the definition of ϕ^E, the vector b is given by $\phi^P(\xi\cdot g)=(\pi(\xi\cdot g),b)$, and the definition of a principal bundle gives $b=\phi_{\pi(\xi)}(\xi\cdot g)=\phi_{\pi(\xi)}(\xi)\cdot g=ag$. The fact that ρ is a homomorphism makes $\rho(ag)\rho(g^{-1})=\rho(a)v$ and ϕ^E is well defined.

Let G be a Lie group, ρ a representation of G on V and M, a manifold. We consider $P=M\times G\xrightarrow{\mathtt{pr}_1}M$, the trivial G-principal bundle on M. Then $E=P\times_\rho V\xrightarrow{p}M$ is **trivial**, i.e. we can build a $\varphi\colon P\times_\rho V\to M\times V$ such that $\mathtt{pr}_1\circ\varphi=p$. It is rather easy: we define

$$\varphi\big[\big((x,g),v\big)\big]=(x,\rho(g)v).$$

It is easy to see that $(\mathtt{pr}_1\circ\varphi)[(x,g),v]=x$ and $p[(x,g),v]=\mathtt{pr}_1(x,g)=x$.

4.5.1 Transition functions

Proposition 4.13.
Let $(\mathcal{U}_\alpha,\phi^P_\alpha)$ be a trivialization of $P\xrightarrow{\pi}M$ whose transition functions are $g_{\alpha\beta}\colon\mathcal{U}_\alpha\cap\mathcal{U}_\beta\to G$. Then $(\mathcal{U}_\alpha,\phi^E_\alpha)$ given by (4.44) is a local trivialization of $E\xrightarrow{p}M$ whose transition functions $g^E_{\alpha\beta}\colon\mathcal{U}_\alpha\cap\mathcal{U}_\beta\to GL(\dim V,\mathbb{K})$ are given by

$$g^E_{\alpha\beta}(x)=\rho(g^P_{\alpha\beta}(x)).$$

Proof. If we write $a:=\phi^E_{\beta x}(\pi^{-1}(x))$, we have $\phi^P_\beta(\pi^{-1}(x))=(x,a)$ and $\phi^E_\alpha\circ(\phi^E_\beta)^{-1}(x,v)=\phi^E_\alpha[(\pi^{-1}(x),\rho(a)^{-1}v)]$. So,

$$\begin{aligned}\phi^E_\alpha[(\pi^{-1}(x),\rho(a)^{-1}v)]&=\Big(x,\rho\big(\phi_{\alpha x}(\pi^{-1}(x))\big)\rho\big(\phi_{\beta x}(\pi^{-1}(x))\big)^{-1}v\Big)\\&=\Big(x,\rho\big(\phi_{\alpha x}(\pi^{-1}(x))\phi_{\beta x}(\pi^{-1}(x))\big)\Big).\end{aligned} \tag{4.45}$$

[4] plongement
[5] Typically $\mathrm{SO}(p,q)$ or $\mathrm{SO}_0(p,q)$.

Then

$$g^E_{\alpha\beta} = \rho\Big(\phi_{\alpha x}(\pi^{-1}(x))\phi_{\beta x}(\pi^{-1}(x))\Big) = \rho(g^P_{\alpha\beta}(x)). \tag{4.46}$$

□

4.5.2 Sections on associated bundle

4.5.2.1 Equivariant functions

We consider a bundle $E = P \times_\rho V \xrightarrow{p} M$ associated with the principal bundle $P \xrightarrow{\pi} M$ and a section $\psi\colon M \to E$.

$$\begin{tikzcd} P \arrow[dr, "\pi^P"'] & & E = P\times_\rho V \arrow[dl, "\pi^E"] \\ & M & \end{tikzcd}$$

A **section** of E is a map $\psi\colon M \to E$ such that $\pi^E \circ \psi = id_M$. We define the function $\hat\psi\colon P \to V$ by

$$\psi(\pi(\xi)) = [\xi, \hat\psi(\xi)]. \tag{4.47}$$

Let us see the condition under which this equation well defines $\hat\psi$. First, remark that a ψ defined by this equation is a section because $p[\xi, v] = \pi(\xi)$, so that $(p\circ\psi)(\pi(\xi)) = \pi(\xi)$. Now, consider a η such that $\pi(\eta) = \pi(\xi)$. Then there exists a $g \in G$ for which $\eta \cdot g = \xi$. For any g and for this one in particular,

$$\psi(\pi(\eta)) = [\eta, \hat\psi(\eta)] = [\eta\cdot g, \rho(g^{-1})\hat\psi(\eta)].$$

Then equation (4.47) defines $\hat\psi$ from ψ if and only if

$$\hat\psi(\xi\cdot g) = \rho(g^{-1})\hat\psi(\xi). \tag{4.48}$$

This condition is called the **equivariance** of $\hat\psi$. Reciprocally, any equivariant function $\hat\psi$ defines a section of $E = P\times_\rho V$.

If $\eta = \xi\cdot g = \chi\cdot k$, one define a sum

$$[\xi, v] + [\chi, w] = [\eta, \rho(g)v + \rho(k)w]. \tag{4.49}$$

If $\psi, \eta\colon M \to E$ are two sections defined by the equivariant functions $\hat\psi, \hat\eta\colon P \to V$, then the section $\psi + \eta$ is defined by the equivariant function $\hat\psi + \hat\eta$.

4.5.2.2 For the endomorphism of sections of E

Let us now make a step backward, and take A in $\operatorname{End}\Gamma(E)$. We will now see that A defines (and is defined by) an equivariant function $\hat A\colon P \to \operatorname{End}V$. Let $\psi\colon M \to E$ be in $\Gamma(E)$. If $\psi(x) = [\xi, v]$, we define the new section $A\psi$ by

$$(A\psi)(x) = [\xi, \hat A(\xi)v] = [\xi, \hat A(\xi)\hat\psi(\xi)].$$

In order for $A\psi$ to be well defined, the function $\hat A$ must satisfy

$$\hat A(\xi\cdot g) = \rho(g^{-1})\hat A(\xi)\rho(g) \tag{4.50}$$

for all g in G.

4.5.2.3 Local expressions

We consider a local trivialization $\phi^P_\alpha\colon \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times G$ of P on $\mathcal{U}_\alpha$ and the corresponding section $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$ given by

$$\sigma_\alpha(x) = (\phi^P_\alpha)^{-1}(x, e).$$

We saw at page 154 that a trivialization of P gives a trivialization of the associated bundle $E = P\times_\rho V$; the definition is

$$\phi^E_\alpha[(\xi, v)] = (\pi(\xi), \rho(a)v) \tag{4.51}$$

if $\phi^P_\alpha(\xi) = (\pi(\xi), a)$. With $\xi = \sigma_\alpha(x)$, we find

$$\phi^E_\alpha[(\sigma_\alpha(x), v)] = (\pi(\sigma_\alpha(x)), \rho(a)v) = (x, v). \tag{4.52}$$

The section ψ can also be seen with respect to the "reference" sections σ_α by means of the definition

$$\psi(x) = [\sigma_\alpha(x), \psi_{(\alpha)}(x)] \tag{4.53}$$

for a function $\psi_{(\alpha)}\colon M \to V$.

Lemma 4.14.
Let $\psi\colon M \to E$ be a section and $\hat{\psi}\colon P \to V$, the corresponding equivariant function. Then

$$\psi_{(\alpha)}(x) = \hat{\psi}(\sigma_\alpha(x)).$$

Proof. By definition, $\psi(x) = \psi(\pi(\xi)) = [\xi, \hat{\psi}(\xi)]$. Thus if we consider in particular $\xi = \sigma_\alpha(x)$,

$$\phi^E_\alpha(\psi(x)) = \phi^E_\alpha[\xi, \hat{\psi}(\xi)] = \phi^E_\alpha[s_\alpha(x), \hat{\psi}(\sigma_\alpha(x))] = (x, \hat{\psi}(\sigma_\alpha(x))). \tag{4.54}$$

□

Let us anticipate. A **spinor** is a section of an associated bundle $E = P \times_\rho V$ where P is a Lorentz-principal bundle, $V = \mathbb{C}^2$ and ρ is the spinor representation of Lorentz on $\mathbb{C}^2$. So a spinor $\psi\colon M \to E$ is *locally* described by a function $\psi_{(\alpha)}\colon M \to \mathbb{C}^2$. The latter is the one that we are used to handle in physics. In this picture, the transformation law of ψ under a Lorentz transformation comes naturally.

Let $\{e_i\}$ be a basis of V; we consider some "reference" sections $\gamma_{\alpha i}$ of the associated bundle $E = P \times_\rho V$ defined by

$$\gamma_{\alpha i}(x) = [\phi_\alpha^{-1}(x, e), e_i]. \tag{4.55}$$

A general section $\psi\colon M \to E$ is defined by an equivariant function $\hat{\psi}\colon P \to V$ which can be written as $\hat{\psi}(\xi) = a^i(\xi)e_i$. If $\eta = \phi_\alpha^{-1}(x, e)$ and $\xi = \eta \cdot g(\xi)$,

$$\psi(x) = [\xi, a^i e_i] = a^i[\eta, \rho(g)e_i] = a^i(\xi)\rho(g(\xi))_i{}^j[\eta, e_j] = c^j(\xi)\gamma_{\alpha j}(x). \tag{4.56}$$

Since the left hand side of this equation just depends on x, the functions c^j must actually not depend on the choice of $\xi \in \pi^{-1}(x)$. So we have $c^j : M \to \mathbb{R}$. Indeed, if we choose $\chi \in \pi^{-1}(x)$,

$$\psi(x) = c^j(\xi)\gamma_{\alpha j}(x) \stackrel{!}{=} [\xi, a^i(\chi)e_i] = \ldots = c^j(\chi)\gamma_{\alpha j}(x),$$

so that $c^j(\xi) = c^j(\chi)$. So any section $\psi\colon M \to E$ can be decomposed (over the open set $\mathcal{U}_\alpha$) as

$$\psi(x) = s^i_\alpha(x)\gamma_{\alpha i}(x). \tag{4.57}$$

4.5.3 Associated and vector bundle

4.5.3.1 General construction

We are going to see that a vector bundle is an associated bundle. For this, we consider a vector bundle $p\colon F \to M$ with a fiber $F_x = V$ of dimension m. Let $G = GL(V)$, P be the trivial principal bundle $P = M \times G$ and ρ be the definition representation of G on V. We set $E = P \times_\rho V$. Our aim is to put a vector bundle structure on E which is equivalent to the one of F. The bijection $b\colon F \to E$ will clearly be

$$b(\phi^{-1}(x, v)) = [(x, e), v]. \tag{4.58}$$

We define the projection $q\colon E \to M$ by

$$q[(x, g), w] = x$$

and we have to show that $q^{-1}(x) = \{[(x, g), w] \text{ st } g \in G \text{ and } w \in V\}$ is a vector space isomorphic to V. The following definitions define a vector space structure:

- multiplication by a scalar: $\lambda[(x, g), v] = [(x, g), \lambda v]$,
- addition: $[(x, g), v] + [(x, h), w] = [(x, e), \rho(g)v + \rho(h)w]$.

As local trivialization map, we consider

$$\begin{aligned} \chi\colon q^{-1}(\mathcal{U}) &\to \mathcal{U} \times V \\ [(x, g), v] &\mapsto (x, \rho(g)v). \end{aligned} \tag{4.59}$$

With this structure, the bijection b is an equivalence because $b|_{F_x}$ is a vector space isomorphism and $q \circ b = p$.

4.5.4 Equivariant functions for a vector field

In order to define in the same way an equivariant function for a vector field $X \in \mathfrak{X}(M)$, we need to see TM as an associated bundle.

Proposition 4.15.
If M is a n dimensional manifold, we have the following isomorphism:

$$\mathrm{SO}(M) \times_{\rho^M} \mathbb{R}^m \simeq TM$$

where $\rho^M\colon \mathrm{SO}(m) \times \mathbb{R}^m \to \mathbb{R}^m$ is defined by $\rho^M(A)v = Av$.

Proof. Recall that an element $b \in \mathrm{SO}(M)_x$ is a map $b\colon \mathbb{R}^m \to T_xM$. The isomorphism is no difficult. It is $\psi\colon \mathrm{SO}(M) \times_{\rho^M} \mathbb{R}^m \to TM$ defined by

$$\psi[b, v] = b(v).$$

It prove no difficult to see that ψ is well defined, injective and surjective. □

Now, let us consider $X \in \mathfrak{X}(M)$. We can see it as an element of $\Gamma(\mathrm{SO}(M) \times_{\rho^M} \mathbb{R}^m)$, and define an equivariant function $\hat{X}\colon \mathrm{SO}(M) \to \mathbb{R}^m$.

Let us make it more explicit. A vector field $Y \in \mathfrak{X}(M)$ is, for each x in M, the data of a tangent vector $Y_x \in T_xM$. Hence the formula $b(v) = Y_x$ defines an element $[b, v]$ in $\mathrm{SO}(M) \times_{\rho^M} \mathbb{R}^m$, and Y defines a section $\tilde{Y}(x) = [b(x), v(x)]$ of $\mathrm{SO}(M) \times_{\rho^M} \mathbb{R}^m$. The associated equivariant function is given by $\hat{Y}(b) = v$ if $b(v) = Y_x$. In other words, the equivariant function $\hat{Y}\colon \mathrm{SO}(M) \to \mathbb{R}^m$ associated with the vector field $Y \in \mathfrak{X}(M)$ is given by

$$\hat{Y}(b) = b^{-1}(Y_x), \tag{4.60}$$

where $x = \pi(b)$.

4.5.5 Gauge transformations

A **gauge transformation** of the G-principal bundle $\pi\colon P \to M$ is a diffeomorphism $\varphi\colon P \to P$ such that

- $\pi \circ \varphi = \pi$,
- $\varphi(\xi \cdot g) = \varphi(\xi) \cdot g$.

When we consider some local sections on $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$, we can describe a gauge transformation with a function $\tilde{\varphi}_\alpha\colon M \to G$ by requiring

$$\varphi(\sigma_\alpha(x)) = \sigma_\alpha(x) \cdot \tilde{\varphi}_\alpha(x).$$

This formula defines φ from $\tilde{\varphi}$ as well as $\tilde{\varphi}$ from φ.

The group of gauge transformations has a natural action on the space of sections given by

$$(\varphi \cdot \psi)(x) = [\varphi(\xi), v]. \tag{4.61a}$$

if $\psi(x) = [\xi, v] = [\xi, \hat{\psi}(\xi)]$. This law can also be seen on the equivariant function $\hat{\psi}$ which defines ψ. The rule is

$$\widehat{\varphi \cdot \psi}(\xi) = \hat{\psi}(\varphi^{-1}(\xi)). \tag{4.61b}$$

Indeed, in the same way as before we find $(\varphi \cdot \psi)(x) = [\xi, \widehat{\varphi \cdot \psi}(x)] \overset{!}{=} [\varphi(\xi), v] = [\varphi(\xi), \hat{\psi}(\xi)]$. Taking $\xi \to \varphi^{-1}(\xi)$ as representative, $(\varphi \cdot \psi)(x) = [\xi, \hat{\psi} \circ \varphi^{-1}(\xi)]$.

4.6 Adjoint bundle

Let $\pi\colon P \to M$ be a G-principal bundle. The **adjoint bundle** is the associated bundle $\mathrm{Ad}(P) = P \times_{\mathrm{Ad}} \mathcal{G}$. An element of that bundle is an equivalent class given by

$$[\xi, X] = [\xi \cdot g, \mathrm{Ad}(g^{-1})X]$$

for every $g \in G$. Here $\xi \in P$ and $X \in \mathcal{G}$.

4.7 Connection on vector bundle: local description

A **connection** on the vector bundle $p\colon E \to M$ is a bilinear map

$$\nabla\colon \mathfrak{X}(M) \times \Gamma(E) \to \Gamma(E),\ (X, s) \quad \mapsto \nabla_X s \tag{4.62}$$

such that

- $\nabla_{fX} s = f\nabla_X s$,
- $\nabla_X(fs) = (X \cdot f)s + f\nabla_X s$

for all $X \in \mathfrak{X}(M)$, $f \in C^\infty(M)$ and $s \in \Gamma(E)$. The operation ∇ is often called a **covariant derivative**.

An easy example is given on the trivial bundle $E = \mathtt{pr}_1\colon M \times \mathbb{C} \to M$. For this bundle, $\Gamma(E) = C^\infty(M, \mathbb{C})$ and the common derivation is a covariant derivation: $\nabla_X s = (ds)X$.

Proposition 4.16.
The value of $(\nabla_X s)(x)$ depends only on X_x and s on a neighbourhood of $x \in M$.

Proof. Let X, $Y \in \mathfrak{X}(M)$ such that $Y_z = f(z)X_z$ with $f(x) = 1$ and $f(z) \neq 1$ everywhere else. Then

$$(\nabla_Y s)(x) - (\nabla_X s)(x) = (f(x) - 1)(\nabla_X s)(x) = 0.$$

Since it is true for any function, the linearity makes that it cannot depend on X_z with $z \neq x$. If we consider now two sections s and s' which are equals on a neighbourhood of x, we can write $s' = fs$ for a certain function f which is 1 on the neighbourhood. Then

$$(\nabla_X s')(x) - (\nabla_X s)(x) = (f(x) - 1)(\nabla_X s)(x) + (Xf)s(x)$$

which zero because on a neighbourhood of x, f is the constant 1. □

This proposition shows that it makes sense to consider only local descriptions of connections. Let $\{e_1, \ldots, e_r\}$ be a basis of V and consider the local sections $\overline{S}_{\alpha i}\colon \mathcal{U}_\alpha \to E$,

$$\overline{S}_{\alpha i}(x) = \phi_\alpha^{-1}(x, e_i).$$

A local section $s_\alpha\colon \mathcal{U}_\alpha \to V$ can be decomposed as $s_\alpha(x) = s_\alpha^i(x)e_i$ with respect to this basis (up to an isomorphism between the different V at each point). Then on $\mathcal{U}_\alpha$,

$$s_\alpha^i \overline{S}_{\alpha i}(x) = s_\alpha^i(x)\phi_\alpha^{-1}(x, e_i) = \phi_\alpha^{-1}(x, s_\alpha^i e_i) = \phi_\alpha^{-1}(x, s_\alpha(x)) = s(x). \tag{4.63}$$

The first equality is the definition of the product $\mathbb{R} \times F \to F$.

So any $s \in \Gamma(E)$ can be (locally !) written under the form[6] $s = s_\alpha^i \overline{S}_{\alpha i}$; in particular $\nabla_X(\overline{S}_{\alpha i})$ can. We define the coefficients θ by

$$\nabla_X(\overline{S}_{\alpha i}) = (\theta_\alpha)_i^j(X)\overline{S}_{\alpha j}. \tag{4.64}$$

where, for each i and j, $(\theta_\alpha)_i^j$ is a 1-form on $\mathcal{U}_\alpha$. We can consider θ_α as a matrix-valued 1-form on $\mathcal{U}_\alpha$.

Proposition 4.17.
The formula

$$(\nabla_X s)_\alpha = Xs_\alpha + \theta_\alpha(X)s_\alpha \tag{4.65}$$

gives a local description of the connection.

Proof. For any $s \in \Gamma(E)$, we have

$$\begin{aligned}
\nabla_X s &= \nabla_X\Big(\sum_j s_\alpha^j \overline{S}_{\alpha j}\Big)\\
&= \sum_j \Big((Xs_\alpha^j)\overline{S}_{\alpha j} + s_\alpha^j \nabla_X \overline{S}_{\alpha j}\Big)\\
&= \sum_i \Big[(Xs_\alpha^i) + s_\alpha^j(\theta_\alpha)_j^i(X)\Big]\overline{S}_{\alpha i}.
\end{aligned}$$

□

[6] be careful on the fact that the "coefficient" s_α^i depends on x : the right way to express this equation is $s(x) = s_\alpha^i(x)\overline{S}_{\alpha i}(x)$.

4.7.1 Connection and transition functions

A connection determines some local matrix-valued 1-forms θ_α on the trivialization $\mathcal{U}_\alpha$. Two natural questions raise. The first is the converse: does a matrix-valued 1-form defines a connection ? The second is to know what is θ_α in function of θ_β on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$? The answer to the latter is given by the following proposition :

Proposition 4.18.
The 1*-form* θ_α *relative to the trivialization* $(\mathcal{U}_\alpha, \phi_\alpha)$ *is related to the* 1*-form* θ_β *relative to the trivialization* $(\mathcal{U}_\beta, \phi_\beta)$ *by*

$$\theta_\beta = g_{\alpha\beta}^{-1} dg_{\alpha\beta} + g_{\alpha\beta}^{-1}\theta_\alpha g_{\alpha\beta}. \tag{4.66}$$

This equation looks like something you know ? If you think to equation (7.3) or (7.11) or any physical equation of gauge transformation for the bosons, then you are almost right.

Proof. We can use equation (4.9) pointwise on $(\nabla_X s)_\alpha$:

$$\begin{aligned}(\nabla_X s)_\alpha &= g_{\alpha\beta}(\nabla_X s)_\beta\\ &= g_{\alpha\beta}\big(Xs_\beta + \theta_\beta(X)s_\beta\big)\\ &= g_{\alpha\beta}\big(X(g_{\alpha\beta}s_\alpha) + \theta_\beta(X)g_{\alpha\beta}s_\alpha\big).\end{aligned} \tag{4.67}$$

We have to compare it with equation (4.65). Note that $g_{\alpha\beta}$ and $\theta_\alpha(X)$ are matrices, then one cannot do

$$g_{\alpha\beta}\theta_\beta(X)g_{\alpha\beta} = g_{\alpha\beta}g_{\alpha\beta}\theta_\beta(X) = \theta_\beta(X)$$

by using $g_{\alpha\beta}g_{\alpha\beta} = \mathbb{1}$. Taking carefully subscripts into account, one sees that the correct form is $(g_{\alpha\beta})^i_j\theta_\beta(X)^j_k(g_{\alpha\beta})^k_l$. Applying Leibnitz formula ($X(fg) = f(Xg) + (Xf)g$), and making the simplification $g_{\alpha\beta}g_{\alpha\beta} = \mathbb{1}$ in the first term, we find

$$\theta_\alpha(X)s_\alpha = g_{\alpha\beta}(Xg_{\alpha\beta})s_\alpha + g_{\alpha\beta}^{-1}\theta_\beta(X)g_{\alpha\beta}s_\alpha.$$

The claim follows from the fact that $Xg_{\alpha\beta} = dg_{\alpha\beta}(X)$. $\square$

Notice that formula (4.66) shows in particular that θ_α takes its values in the Lie algebra $\mathfrak{gl}(V)$, see for example subsection 2.2.2.
The inverse is given in the

Proposition 4.19.
If we choose a family of $\mathfrak{gl}(V)$*-valued* 1*-forms* θ_α *on* $\mathcal{U}_\alpha$ *satisfying* (4.66)*,then the formula*

$$(\nabla_X s)_\alpha = Xs_\alpha + \theta_\alpha(X)s_\alpha$$

defines a connection on E*.*

Proof. Note that θ is $C^\infty(M)$-linear, thus

$$(\nabla_{fX} s)_\alpha = (fX)s_\alpha + \theta_\alpha(fX)s_\alpha = f[Xs_\alpha + \theta_\alpha(X)s_\alpha] = f(\nabla_X s)_\alpha. \tag{4.68}$$

In expressions such that $\theta_\alpha(X)(fs_\alpha)$, the product is a matrix times vector product between $\theta_\alpha(X)$ and s_α; the position of the f is not important. So we can check the second condition :

$$\begin{aligned}(\nabla_X(fs))_\alpha &= X(fs_\alpha) + \theta_\alpha(X)(fs_\alpha)\\ &= X(f)s_\alpha + f(Xs_\alpha) + f\theta_\alpha(X)s_\alpha\\ &= df(X)s_\alpha + f(\nabla_X s)_\alpha.\end{aligned} \tag{4.69}$$

This concludes the proof. $\square$

4.7.2 Torsion and curvature

The map $T^\nabla \colon \mathfrak{X}(X) \times \mathfrak{X}(X) \to \mathfrak{X}(X)$ defined by

$$T^\nabla(X,Y) = \nabla_X Y - \nabla_Y X - [X,Y] \tag{4.70}$$

is the **torsion** of the connection ∇. When $T^\nabla(X,Y) = 0$ for every X and Y in $\mathfrak{X}(X)$, we say that ∇ is a **torsion free** connection. Let X, Y be in $\mathfrak{X}(M)$, and consider the map $R(X,Y)\colon \Gamma(E) \to \Gamma(E)$ defined by

$$R(X,Y)\colon \Gamma(E) \to \Gamma(E)\; s \quad \mapsto \nabla_X\nabla_Y s - \nabla_Y\nabla_X s - \nabla_{[X,Y]}s. \tag{4.71}$$

For each $x \in M$, R can be seen as a bilinear map $R\colon T_xM \times T_xM \to \mathrm{End}(E_x)$. It is called the **curvature** of the connection ∇. For every $f \in C^\infty(M)$, it satisfies

$$R(fX, Y)s = fR(X, Y)s = R(X, Y)fs.$$

In a trivialization $(\mathcal{U}_\alpha, \phi_\alpha)$, we have $(\nabla_X s)_\alpha = Xs_\alpha + \theta_\alpha(X)s_\alpha$. In the expression of $(R(X,Y)s)_\alpha$, the terms coming from the Xs_α part of covariant derivative make

$$XYs_\alpha - YXs_\alpha - [X,Y]s_\alpha = 0.$$

The other terms are no more than matricial product, hence the formula

$$(R(X,Y)s)_\alpha = \Omega_\alpha(X,Y)s_\alpha \tag{4.72}$$

defines a 2-form Ω_α which takes values in $GL(r, \mathbb{K})$. We can find an expression for Ω in terms of θ :

$$\Omega_\alpha(X,Y) = X\theta_\alpha(Y) - Y\theta_\alpha(X) - \theta_\alpha([X,Y]) + \theta_\alpha(X)\theta_\alpha(Y) - \theta_\alpha(Y)\theta_\alpha(X);$$

it is written as

$$\Omega_\alpha = d\theta_\alpha + \theta_\alpha \wedge \theta_\alpha = d\theta_\alpha + \frac{1}{2}[\theta_\alpha, \theta_\alpha] \tag{4.73}$$

which is a notational shortcut for

$$\Omega_\alpha(X,Y) = d\theta_\alpha(X,Y) + [\theta_\alpha(X), \theta_\alpha(Y)]. \tag{4.74}$$

These equations are called **structure equations**. Pointwise, the second term is a matrix commutator; be careful on the fact that, when we will speak about principal bundle, the forms θ's will take their values in a Lie algebra. On $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$, we have

$$\Omega_\beta(X,Y) = g_{\alpha\beta}^{-1}\Omega_\alpha(X,Y)g_{\alpha\beta}.$$

The curvature and the connection fulfill the **Bianchi identities** :

Lemma 4.20.

$$d\Omega_\alpha + [\theta_\alpha \wedge \Omega_\alpha] = 0.$$

Proof. For each matricial entry, θ_α is a 1-form on $\mathcal{U}_\alpha$, then $\theta_\alpha(X)$ is a function which to $x \in M$ assign $\theta_\alpha(x)(X_x) \in \mathbb{R}$. So we can apply d and Leibnitz on the product $\theta_\alpha(X)\theta_\alpha(Y)$.

$$d\big(\theta_\alpha(X)\theta_\alpha(Y)\big) = \theta_\alpha(X)d\theta_\alpha(Y) + d\theta_\alpha(X)\theta_\alpha(Y).$$

Differentiating equation (4.73), $d\Omega_\alpha = d\theta_\alpha \wedge \theta_\alpha - \theta_\alpha \wedge d\theta_\alpha$. □

4.7.3 Divergence, gradient and Laplacian

We define the **gradient** of a function $f \in C^\infty(M)$, denoted by ∇f as the vector field such that

$$g(\nabla f, X) = X(f). \tag{4.75}$$

The **divergence** of a vector field $X \in \Gamma(TM)$, is the function $\nabla \cdot X \in C^\infty(M)$ defined by

$$(\nabla \cdot X)(x) = \mathrm{Tr}\left(v \mapsto \nabla_v X\right) \tag{4.76}$$

where the trace is the one of $v \mapsto \nabla_v X$ seen as an operator on T_xM. The **Laplacian** of the function f is the function Δf given by

$$\Delta f = \nabla \cdot (\nabla f). \tag{4.77}$$

4.8 Connexion on vector bundle: algebraic view

A **connection** on the vector bundle $\pi\colon E \to M$ is a linear map

$$\nabla\colon \Gamma^\infty(E) \to \Gamma^\infty(E) \otimes \Omega^1(M)$$

which satisfies the Leibnitz rule

$$\nabla(\sigma f) = (\nabla\sigma)f + \sigma \otimes df \tag{4.78}$$

for any section $\sigma\colon M\to E$ and function $f\colon M\to\mathbb{C}$. If $\{\sigma_i\}$ is a local basis of E, one can write $\sigma=\sigma_i f^i$ and one defines the **Christoffel symbols** $\Gamma^j_{i\mu}$ in this basis by

$$\nabla\sigma=\nabla(\sigma_i f^i)=(\nabla\sigma_i)f^i+\sigma_i\otimes f(f^i)=f^i\Gamma^j_{i\mu}\sigma_j\otimes dx^\mu+\sigma_i\otimes d(f^i). \tag{4.79}$$

The notations $d\sigma=\sigma_i\otimes d(f^i)$ and $\Gamma\sigma=f^i\Gamma^j_{i\mu}\sigma_j\otimes dx^\mu$ lead us to the compact usual form

$$\nabla\sigma=(d+\Gamma)\sigma.$$

When $E=TM$ over a (pseudo)Riemannian manifold M, we know the Levi-Civita connection which is compatible with the metric:

$$g(\nabla X,Y)+g(X,\nabla Y)=d\big(g(X,Y)\big). \tag{4.80}$$

One can see g as acting on $\big(\mathfrak{X}(M)\otimes\Omega^1(M)\big)\times\mathfrak{X}(M)$ with

$$g\big(r^i_\nu\partial_i\otimes dx^\nu,t^j\partial_j\big):=r^i_\nu j^j g(\partial_i,\partial_j)dx^\nu,$$

which at each point is a form. From condition (4.80), we see ∇ as a Levi-Civita connection on the bundle $E=T^*M$ which values in

$$\Gamma^\infty(T^*M)\otimes\Omega^1(M)\simeq\Omega^1(M)\otimes\Omega^1(M).$$

This is defined as follows. A 1-form ω can always be written under the for $\omega=X^\flat:=g(X,.)$ for a certain $X\in\mathfrak{X}(M)$. Then (4.80) gives

$$(\nabla X)^\flat Y+\omega(\nabla Y)=d(\omega Y),$$

and we put $\nabla\omega=(\nabla X)^\flat$, i.e

$$(\nabla\omega)Y=d(\omega Y)-\omega(\nabla Y) \tag{4.81}$$

for all $Y\in\mathfrak{X}(M)$. When $\omega=dx^i$ and $Y=\partial_j$, we find

$$(\nabla dx^i)\partial_j=d(dx^i\partial_j)-dx^i(\nabla\partial_j)=d(\delta^i_j)-\Gamma^l_{jk}\partial_l\otimes dx^k=-\Gamma^l_{jk}\delta^i_l\otimes dx^k=-\Gamma^i_{jk}\,dx^k. \tag{4.82}$$

So we get the local formula

$$\nabla dx^i=-\Gamma^i_{jk}\,dx^j\otimes dx^k. \tag{4.83}$$

If the form writes locally $\omega=dx^i f_i$,

$$\nabla\omega=\nabla(dx^i)f_i+dx^i\otimes df_i=-f_i\Gamma^i_{jk}\,dx^j\otimes dx^k+d\omega=(d-\tilde\Gamma)\omega \tag{4.84}$$

where we taken the notations $d\omega=dx^i\otimes df_i$ and $\tilde\Gamma\omega=f_i\Gamma^i_{jk}dx^j\otimes dx^k$.

4.8.1 Exterior derivative

If E is a m-dimensional vector bundle over M and $s\colon M\to E$ is a section, we say that a **exterior derivative** is a map $D\colon\Gamma(E)\to\Gamma(E\otimes\Omega^1M)$ such that for every $f\in C^\infty(M)$ we have

$$D(fs)=s\otimes df+f(Ds).$$

An exterior derivative can be extended to $D\colon\Gamma(E\otimes\Omega^kM)\to\Gamma(E\otimes\Omega^{k+1}M)$ imposing the condition

$$D(\omega\wedge\alpha)=(D\omega)\wedge\alpha+(-1)^k\omega\wedge d\alpha \tag{4.85}$$

for every $\omega\in\Gamma(E\otimes\Omega^kM)$ and $\alpha\in\Gamma(E\otimes\Omega^lM)$. The result is an element of $\Gamma(E\otimes\Omega^{k+l+1}M)$.

Coordinatewise expressions are obtained when one choose a specific section (e_i) of the frame bundle of E. In that case for each i, the derivative e_i is an element of $\Gamma(E\otimes\Omega^1M)$ and we define $\omega^j_i\in\Omega^1(M)$ by

$$De_i=\sum_{j=1}^k e_j\otimes\omega^j_i. \tag{4.86}$$

For each i and j, we have an element $\omega^j_i\in\Omega^1(M)$, so that we say that $\omega\in\Omega^1(M,\mathfrak{gl}(m))$. Now a section can be expressed as $s=s^ie_i$ where s^i are functions, so we have

$$D(s)=D(s^ie_i)=e_i\otimes ds^i+s^iD(e_i)=e_i\otimes ds^i+s^ie_j\otimes\omega^j_i=e_i\otimes ds^i+e_i\otimes s^j\omega^i_j. \tag{4.87}$$

Expressed in component, we find $D(s)^i=ds^i+s^j\omega^i_j$, so that we often write

$$D=d+\omega. \tag{4.88}$$

When a section e is given, we write $s = s^i(e)e_i$, indicating the dependence of the functions s^i in the choice of the frame e :

$$D(s) = e_i \otimes ds^i(e) + e_i \otimes s^j(e)\omega(e)^i_j.$$

When we apply both sides to a vector $X \in \Gamma(TM)$, we find

$$D_X(s) = e_i \otimes \Big(X(s^i) + s^j\omega^i_j(X)\Big). \tag{4.89}$$

By convention we say that, when $f \in C^\infty(M)$, is a function, D_X reduces to the action of the vector field X:

$$D_X(f) = X(f). \tag{4.90}$$

4.8.1.1 Covariant exterior derivative

An important exterior derivative is the covariant exterior derivative. If the vector bundle E is endowed by a covariant derivative ∇, we define the corresponding **covariant exterior derivative** by the following :

1. for a section $s\colon M \to E$ (i.e. a 0-form) we define

$$(d_\nabla s)(X) = \nabla_X s, \tag{4.91}$$

2. and on the 1-form $\sum_i\big(s_i \otimes \omega_i\big) \in \Gamma(E \otimes T^*M)$,

$$d_\nabla\big(\sum_i s_i \otimes \omega_i\big) = \sum_i(d_\nabla s_i) \wedge \omega_i + \sum_i s_i \otimes d\omega_i. \tag{4.92}$$

The latter relation is the condition (4.85) with $k = 0$.

4.8.1.2 Soldering form and torsion

Let us particularize to the case where E has the same dimension as the manifold. In that case, we can introduce a **soldering form**, that is an element $\theta \in \Omega^1(M, E)$ such that for every $x \in M$ the map $\theta_x\colon T_xM \to E_x$ is a vector space isomorphism. When a soldering form θ is given, the **torsion** is the exterior derivative D is

$$T = D\theta. \tag{4.93}$$

Using a local frame e, we have forms $\theta^i(e) \in \Omega^1(M)$ such that

$$\theta(X) = \theta^i(X)e_i.$$

We see θ as an element of $\Gamma(E \otimes \Omega^1(M))$ by identifying $\theta = e_i \otimes \theta^i$. Thus we have

$$D\theta = D(e_i \otimes \theta^i) = De_i \wedge \theta^i(e) + e_i \wedge d\theta^i(e) = (e_j \otimes^j_i) \wedge \theta^i(e) + e_i \wedge d\theta^i(e),$$

or in coordinates :

$$(D\theta)^i = \omega^i_j \wedge \theta^j(e) + d\theta^i(e). \tag{4.94}$$

Notice that it provides the formula

$$T = d_\omega\theta \tag{4.95}$$

for the torsion as exterior covariant derivative of the connection form.

4.8.1.3 Example : Levi-Civita

We consider the vector bundle $E = TM$ and the local basis $e_i = \partial_i$. An exterior derivative in this case is a map $D\colon \Gamma(TM) \to \Gamma\Big(TM \otimes \Omega^1 M\Big)$. In that particular case, we denote by $\nabla_X Y$ the vector field $D(Y)X$, and it is computed by first writing $D(X)_x = \sum_i Z^i_x \otimes \omega^i_x$ with $Z^i \in \Gamma(TM)$ and $\omega^i \in \Omega^1(M)$. The we have

$$D(X)_x Y_x = \omega + x^i(Y_x)Z^i_x. \tag{4.96}$$

A good choice of soldering form is $\theta_x = \mathrm{id}$ for every $x \in M$, or $\theta(X) = X$. In coordinates, that soldering form is given by $\theta^i(\partial_j) = \delta^i_j$. The **Christoffel symbols** are defined by

$$\nabla_{\partial_i}\partial_j = \Gamma^k_{ij}\partial_k, \tag{4.97}$$

and the covariant derivative reads

$$\nabla_X Y = \nabla_{X^i\partial_i}(Y^j\partial_j) = X^i\Big((\partial_i Y^j)\partial_j + Y^j\nabla_{\partial_i}\partial_j\Big) = \Big(X(Y^k) + X^iY^j\Gamma^k_{ij}\Big)\partial_k. \tag{4.98}$$

We can determine the Christoffel symbols in function of the connection form using the fact that on the one hand, $\nabla_{\partial_i}\partial_j = \Gamma^k_{ij}\partial_k$, and on the other hand,

$$\nabla_{\partial_i}\partial_j = D(\partial_j)(\partial_i) = \partial_k \otimes \omega^k_j(\omega_i),$$

so that

$$\Gamma^k_{ij} = \omega^k_j(\partial_i) \tag{4.99}$$

Now we can get the same result as equation (4.98) using the exterior derivative formalism. First we have $DY = \partial_i \otimes dY^i + \partial_i \otimes X^j\omega^i_j$, so that

$$(DY)X = \partial_i \otimes dY^i(X) = \partial_i \otimes X^j\omega^i_j(X^k\partial_k),$$

in which we use the relation $\omega^i_j(X^k\partial_k) = X^k\omega^i_j(\partial_k) = X^k\Gamma^i_{jk}$ to get

$$(DY)X = \big(X(Y^i) + X^jX^k\Gamma^i_{jk}\big)\partial_i.$$

Notice that the anti-symmetric part of Γ with respect to its two lower indices does not influence the covariant derivative. Let us compute the torsion in terms of Γ. For that remark that $d\theta^i = 0$ because

$$(d\theta^i)(X,Y) = X\theta^i(Y) - Y\theta^i(X) - \theta^i\big([X,Y]\big) = X(Y^i) - Y(X^i) - [X,Y]^i = 0.$$

Thus we have

$$\begin{aligned}(D\theta)(\partial_k \otimes \partial_l) &= \big((D\partial_i)\partial_k\big)\theta^i(\partial_l) - \big((D\partial_i)\partial_l\big)\theta^i(\partial_k)\\ &= \delta^i_l\Gamma^j_{ik}\partial_j - \delta^i_k\Gamma^j_{il}\partial_j\\ &= (\Gamma^j_{lk} - \Gamma^j_{kl})\partial_j.\end{aligned}$$

The connection ∇ is moreover compatible with the metric because

$$\nabla_Z\big(g(X,Y)\big) = Z\big(\eta(eX,eY)\big) = \eta\big(\underbrace{D_Z(eX)}_{=e(\nabla_Z X)}, eY\big) + \eta\big(eX, D_Z(eY)\big) = g(\nabla_Z X, Y) + g(X, \nabla_Z Y).$$

4.9 Connection on principal bundle

4.9.1 First definition: 1-form

We consider a G-principal bundle

$$\begin{array}{ccc} G & \rightsquigarrow & P \\ & & \downarrow \pi \\ & & M \end{array}$$

and $\mathcal{G}$, the Lie algebra of G.

Definition 4.21.
*A **connection** on P is a 1-form $\omega \in \Omega(P,\mathcal{G})$ which fulfills*

- $\omega_\xi(A^*_\xi) = A,$
- $(R^*_g\omega)_\xi(\Sigma) = \mathrm{Ad}(g^{-1})(\omega_\xi(\Sigma)),$

for all $A \in \mathcal{G}$, $g \in G$, $\xi \in P$ and $\Sigma \in T_\xi P$

Here, R_g is the right action: $R_g\xi = \xi \cdot g$ and A^* stands for the **fundamental field** associated with A for the action of G on P:

$$A^*_\xi = \frac{d}{dt}\Big[\xi \cdot e^{-tA}\Big]_{t=0}, \tag{4.100}$$

For each $\xi \in P$, we have $\omega_\xi \colon T_\xi P \to \mathcal{G}$. See section 2.3.

If α is a connection 1-form on P, we say that Σ is an **horizontal** vector field if $\alpha_\xi(\Sigma) = 0$ for all $\xi \in P$. If $X_x \in T_xM$ and $\xi \in \pi^{-1}(x)$, there exists an unique[7] Σ in $T_\xi P$ which is horizontal and such that $\pi_*(\Sigma) = X_x$. This Σ is called the **horizontal lift** of X_x. We can also pointwise construct the horizontal lift of a vector field. The one of X is often denoted by $\overline{X}$; it is an element of $\mathfrak{X}(P)$.

[7] See [1], chapter II, proposition 1.2.

4.9.2 Second definition: horizontal space

For each $\xi \in P$, we define the **vertical space** $V_\xi P$ as the subspace of $T_\xi P$ whose vectors are tangent to the fibers: each $v \in V_\xi P$ fulfills $d\pi v = 0$. Any such vector is given by a path contained in the fiber of ξ. So, $v \in V_\xi P$ if and only if there exists a path $g(t) \in G$ such that $v = \frac{d}{dt}\Big[\xi \cdot g(t)\Big]_{t=0}$.

A **connection** Γ is a choice, for each $\xi \in P$, of an **horizontal space** $H_\xi P$ such that

- $T_\xi P = V_\xi P \oplus H_\xi P$,
- $H_{\xi\cdot g} = (dR_g)_\xi H_\xi$,
- $H_\xi P$ depends on ξ under a differentiable way.

The second condition means that the distribution $\xi \to H_\xi$ is invariant under G. Thanks to the first one, for each $X \in T_\xi P$, there exists only one choice of $Y \in H_\xi P$ and $Z \in V_\xi P$ such that $X = Y + Z$. These are denoted by vX and hX and are naturally named *horizontal* and *vertical components* of X. The third condition means that if X is a differentiable vector field on P, then vX and hX are also differentiable vector fields. We will often write V_ξ and H_ξ instead of $V_\xi P$ and $V_\xi P$.

The word *connection* probably comes from the fact that the horizontal space gives a way to jump from a fiber to the next one. When we consider a connection Γ, we can define a $\mathcal{G}$-valued connection 1-form by

$$\omega(X)^*_\xi = vX_\xi.$$

The existence is explained in section 2.3. It is clear that $\omega(X) = 0$ if and only if X is horizontal. The theorem which connects the two definitions is the following.

Theorem 4.22.
If Γ is a connection on a G-principal bundle, and ω is its 1-form, then

(i) for any $A \in \mathcal{G}$, we have $\omega(A^*) = A$,

(ii) $(R_g)^*\omega = \mathrm{Ad}(g^{-1})\omega$, i.e. for any $X \in T_\xi P$, $g \in G$ and $\xi \in M$,

$$\omega((dR_g)_\xi X) = \mathrm{Ad}(g^{-1})\omega_\xi(X)$$

Conversely, if one has a $\mathcal{G}$-valued 1-form on P which fulfills these two requirement, then one has one and only one connection on P whose associated 1-form is ω.

Proof. (i) The definition of ω is $\omega(X)^*_\xi = vX$. Then $\omega(A^*)^*_\xi = vA^*_\xi = A^*_\xi$ because A^* is vertical. From lemma 2.10, $\omega(A^*) = A$.

(ii) Let $X \in \mathfrak{X}(P)$. If X is horizontal, the definition of a connection makes dR_dX also horizontal, then the claim becomes $0 = 0$ which is true. If X is vertical, there exists a $A \in \mathcal{G}$ such that $X = A^*$ and a lemma shows that dR_gX is then the fundamental field of $\mathrm{Ad}(g^{-1})A$. Using the properties of a connection,

$$(R_g^*\omega)_\xi(X) = \omega_{\xi\cdot g}(dR_gX) = \mathrm{Ad}(g^{-1})A = \mathrm{Ad}(g^{-1})\omega_\xi(X). \tag{4.101}$$

Now we turn our attention to the inverse sense: we consider a 1-form which fulfills the two conditions and we define

$$H_\xi = \{X \in T_\xi P \text{ st } \omega(X) = 0\}. \tag{4.102}$$

We are going to show that this prescription is a connection. First consider a $X \in V_\xi$, then $X = A^*$ and $\omega(X) = A$. So $H_\xi \cap V_\xi = 0$. Now we consider $X \in T_\xi P$ and we decompose it as

$$X = A^* + (X - A^*)$$

where A^* is the vertical component of X. If $\omega(dR_gX) = 0$ for all $g \in G$, then $\omega(X) = 0$, then a vector $X \in H_\xi$ fulfills at most $\dim G$ independent constraints $\omega(dR_gX) = 0$ and $\dim H_\xi$ is at least $\dim P - \dim G$. On the other hand, $\dim V_\xi = \dim G$; then

$$\dim V_\xi + \dim H_\xi \geqslant \dim G + \dim P - \dim G.$$

Then the equality must holds and $V_\xi \oplus H_\xi = T_\xi P$.

We have now to prove that ω is the connection form of H_ξ, i.e. that $\omega(X)$ is the unique $A \in \mathcal{G}$ such that A^*_ξ is the vertical component of X. Indeed if $X \in T_\xi P$, it can be decomposed as into $A^* \in V_\xi$ and $Y \in H_\xi$ and

$$\omega(X) = \omega(A^* + Y) = \omega(A^*) = A.$$

It remains to be proved that the horizontal space H_ξ of any connection Γ is related to the corresponding 1-form ω by $H_\xi = \{X \in T_\xi P \text{ st } \omega_\xi(X) = 0\}$. From the connection Γ, the 1-form is defined by the requirement that $\omega(X)^*_\xi = vX_\xi$. For $X \in H_\xi$, it is clear that $vX = 0$, so that $\omega(X)^* = 0$. This implies $\omega(X) = 0$ because we suppose that the action of G is effective.

□

The projection $\pi\colon P \to M$ induces a linear map $d\pi\colon T_\xi P \to T_x M$. We will see that, when a connection is given, it is an isomorphism between H_ξ and $T_x M$ (if $x = \pi(\xi)$). The **horizontal lift** of $X \in \mathfrak{X}(M)$ is the unique horizontal vector field (i.e. it is pointwise horizontal) such that $d\pi(\overline{X}_\xi) = X_{\pi(\xi)}$. The proposition which allows this definition is the following.

Proposition 4.23.
For a given connection on the G-principal bundle P and a vector field X on M, there exists an unique horizontal lift of X. Moreover, for any $g \in G$, the horizontal lift is invariant under dR_g.

The inverse implication is also true: any horizontal field on P which is invariant under dR_g for all g is the horizontal lift of a vector field on M.

This proposition comes from [1], chapter II, proposition 1.2.

Proof. We consider the restriction $d\pi\colon H_\xi \to T_{\pi(\xi)}M$. It is injective because $d\pi(X - Y)$ vanishes only when $X - Y$ is vertical or zero. Then it is zero. It is cleat that $d\pi\colon T_\xi P \to T_{\pi(\xi)}M$ is surjective. But $d\pi X = 0$ if X is vertical, then $d\pi$ is surjective from only H_ξ.

So we have existence and unicity of an horizontal lift. Now we turn our attention to the invariance. The vector $dR_g\overline{X}_\xi$ is a vector at $\xi \cdot g$. From the definition of a connection, $dR_g H_\xi = H_{\xi\cdot g}$, then $dR_g\overline{X}_\xi$ is the unique horizontal vector at $\xi \cdot g$ which is sent to X_x by $d\pi$. Thus it is $\overline{X}_{\xi\cdot g}$.

For the inverse sense, we consider $\overline{X}$, an horizontal invariant vector field on P. If $x \in M$, we choose $\xi \in \pi^{-1}(x)$ and we define $X_x = d\pi(\overline{X}_\xi)$. This construction is independent of the choice of ξ because for $\xi' = \xi \cdot g$, we have

$$d\pi(\overline{X}_{\xi'}) = \pi(dR_g\overline{X}_\xi) = \pi(\overline{X}_\xi).$$

□

An other way to see the invariance is the following formula:

$$\overline{X}_{\xi\cdot g} = (dR_g)_\xi \overline{X}_\xi.$$

By definition, $\overline{X}_{\xi\cdot g}$ is the unique vector of $T_{\xi\cdot g}P$ which fulfils $d\pi\overline{X}_{\xi\cdot g} = X_x$ if $\xi\pi^{-1}(x)$, so the following computation proves the formula:

$$(d\pi)_{\xi\cdot g}((dR_g)_\xi\overline{X}_\xi) = d(\pi \circ R_g)_\xi\overline{X}_\xi = d\pi_\xi\overline{X}_\xi = X_x. \tag{4.103}$$

4.9.3 Curvature

The curvature of a vector or associated bundle satisfies $\Omega_\alpha = d\theta_\alpha + \theta_\alpha \wedge \theta_\alpha$. So we naturally define the **curvature** of the connection ω on a principal bundle as the $\mathcal{G}$-valued 2-form

$$\Omega = d\omega + \omega \wedge \omega. \tag{4.104}$$

When we consider a local section $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$ on $\mathcal{U}_\alpha \subset M$, we can express the curvature with a 2-form on M instead of P by the formula

$$F_{(\alpha)} = \sigma^*_\alpha\Omega,$$

or, more explicitly, by $F_{(\alpha)x}(X,Y) = \Omega_{\sigma_\alpha(x)}(d\sigma_\alpha X, d\sigma_\alpha Y))$. Note that if $\mathcal{G}$ is abelian, $\Omega = d\omega$ and $d\Omega = 0$.

4.10 Exterior covariant derivative and Bianchi identity

Let $\omega \in \Omega^1(P,\mathcal{G})$ be a connection 1-form on the G-principal bundle P. Using the operation $[. \wedge .]$ defined in section 4.3, we define the **exterior covariant derivative** by

$$d_\omega\alpha = d\alpha + \frac{1}{2}[\omega \wedge \alpha] \qquad \text{when } \alpha \in \Omega^1(P,\mathcal{G}), \tag{4.105}$$

$$d_\omega\beta = d\beta + [\omega \wedge \beta] \qquad \text{when } \beta \in \Omega^2(P,\mathcal{G}), \tag{4.106}$$

The **curvature** is the 2-form defined by

$$\Omega = d_\omega \omega = d\omega + \omega \wedge \omega \tag{4.107}$$

where d_ω is the exterior covariant derivative associated with the connection form ω, and the wedge has to be understood as in equation (4.12).

Proposition 4.24.
The curvature form satisfies the identity

$$d_\omega \Omega = 0 \tag{4.108}$$

which is the Bianchi identity

Proof. taking the differential of $\Omega = d\omega + \omega \wedge \omega$, we find

$$d\Omega = d^2\omega + d\omega \wedge \omega - \omega \wedge d\omega$$

in which $d^2\omega = 0$ and we replace $d\omega$ by $\Omega - \omega \wedge \omega$, so that

$$d\Omega = \Omega \wedge \omega - \omega \wedge \Omega,$$

which becomes the Bianchi identity using the definition of d_ω and the notation (4.14). □

Remark that the Bianchi identity reads $d_\omega^2 \omega = 0$, but that in general d_ω does not square to zero.

4.11 Covariant derivative on associated bundle

Now we consider a general G-principal bundle $\pi\colon P \to M$ and an associated bundle $E = P \times_\rho V$. We define a product $\mathbb{R} \times E \to E$ by

$$\lambda[\xi, v] = [\xi, \lambda v]. \tag{4.109}$$

It is clear that the equivariant function $\widehat{\lambda\psi}$ defines the section $\lambda\psi$. A **covariant derivative** is a map

$$\nabla\colon \mathfrak{X}(M) \times \Gamma(M, E) \to \Gamma(M, E)\ (X, \psi) \quad \mapsto \nabla_X \psi \tag{4.110}$$

such that

$$\nabla_{fX}\psi = f\nabla_X\psi, \tag{4.111a}$$

$$\nabla_X(f\psi) = (X \cdot f)\psi + f\nabla_X\psi \tag{4.111b}$$

where products have to be understood by formula (4.109).

Theorem 4.25.
A connection on a principal bundle gives rise to a covariant derivative on any associated bundle by the formula

$$\widehat{\nabla^E_X \psi}(\xi) = \overline{X}_\xi(\hat{\psi}) \tag{4.112}$$

where $\hat{\psi}\colon P \to V$ is the function associated with the section $\psi\colon M \to E$.

We have to prove that it is a good definition: the function $\widehat{\nabla^E_X \psi}$ must define a section $\nabla^R_X \psi\colon M \to E$ and the association $\psi \to \nabla^E_X \psi$ must be a covariant derivative.

With the discussion of page 14 about the application of a tangent vector on a map between manifolds, we have $(d\varphi X)f = X(f \circ \varphi)$. By using this equality in the case of $\overline{X}$ with $\hat{\psi}$ and R_g, we find $(dR_g\overline{X})(\hat{\psi}) = \overline{X}(\hat{\psi} \circ R_g)$ and thus

$$\overline{X}_{\xi \cdot g}(\hat{\psi}) = \overline{X}_\xi(dR_g \hat{\psi}).$$

We prove the theorem step by step.

Proposition 4.26.
The function $\widehat{\nabla^E_X \psi}$ defines a section of P.

Proof. We have to see that $\widehat{\nabla^E_X \psi}$ is an equivariant function. The equivariance of $\hat{\psi}$ gives $\hat{\psi} \circ R_g = \rho(g^{-1})\hat{\psi}$, thus

$$\widehat{\nabla^E_X \psi}(\xi \cdot g) = \overline{X}_{\xi \cdot g}(\hat{\psi}) = \big((dR_g)_\xi \overline{X}_\xi\big)(\hat{\psi}) = \overline{X}_\xi(\hat{\psi} \circ R_g) = \overline{X}_\xi(\rho(g^{-1})\hat{\psi}) = \rho(g^{-1})\overline{X}_\xi(\hat{\psi}). \tag{4.113}$$

The last equality comes from the fact that the product $\rho(g^{-1})\hat{\psi}$ is a linear product "matrix times vector" and that $\overline{X}_\xi$ is linear. □

Theorem 4.27.
The definition

$$\widehat{\nabla^E_X\psi}(\xi) = \overline{X}_\xi(\hat{\psi})$$

defines a covariant derivative.

Proof. We have to check the two conditions given on page 158.
First condition. By definition, $\widehat{\nabla^E_{fX}\psi}(\xi) = \overline{fX}_\xi(\hat{\psi})$. Now we prove that

$$\overline{fX}_\xi(\hat{\psi}) = (f \circ \pi)(\xi)\overline{X}_\xi(\hat{\psi}). \tag{4.114}$$

This formula is coherent because $\overline{X}_\xi(\hat{\psi}) \in V$ and $(f \circ \pi)(\xi) \in \mathbb{R}$. By definition of the horizontal lift, $\overline{fX}_\xi$ is the unique vector such that

- $d\pi_\xi(\overline{fX}_\xi) = (fX)_x = f(x)d\pi\overline{X}_\xi = (f \circ \pi)(\xi)d\pi\overline{X}_\xi$,
- $\omega_\xi(\overline{fX}_\xi) = 0$.

We check that $(f \circ \pi)(\xi)\overline{X}_\xi$ also fulfills these two conditions because $d\pi$ and ω are $C^\infty(P)$-linear. Equation (4.114) immediately gives

$$\widehat{\nabla^E_{fX}\psi}(\xi) = (f \circ \pi)(\xi)\widehat{\nabla^E_X\psi}(\xi). \tag{4.115}$$

Now we show that $\widehat{f\nabla^E_X\psi}$ is the same. The section $f\nabla^E_X\psi\colon M \to E$ is given by $(f\nabla^E_X\psi)(x) = f(x)(\nabla^E_X\psi)(x)$, and by definition of the associated equivariant function,

$$f(x)(\nabla^E_X\psi)(x) = [\xi, f(x)\widehat{\nabla^E_X\psi}(\xi)].$$

Then

$$\widehat{f\nabla^E_X\psi}(\xi) = f(x)\widehat{\nabla^E_X\psi}(\xi) = (f \circ \pi)(\xi)\widehat{\nabla^E_X\psi}(\xi). \tag{4.116}$$

All this shows that $\nabla^E_{fX}\psi = f\nabla^E_X\psi$.
Second condition. This is a computation using the Leibniz rule:

$$\begin{aligned}
\widehat{\nabla^E_X(f\psi)}(\xi) &= \overline{X}_\xi(\widehat{f\psi}) \overset{(a)}{=} \overline{X}_\xi((\pi \circ f)\hat{\psi}) \\
&\overset{(b)}{=} \overline{X}_\xi(\pi^* f)\hat{\psi}(\xi) + (\pi^* f)(\xi)\overline{X}_\xi\hat{\psi} = d(f \circ \pi)_\xi\overline{X}_\xi\hat{\psi}(\xi) + f\widehat{\nabla^E_X\psi}(x) \\
&= df_{\pi(\xi)}d\pi_\xi\overline{X}_\xi\hat{\psi}(\xi) + f\widehat{\nabla^E_X\psi}(x) = X_x(f)\hat{\psi}(\xi) + f\widehat{\nabla^E_X\psi}(x) \\
&= \widehat{(Xf)\psi}(\xi) + \widehat{f\nabla^E_X\psi}(\xi)
\end{aligned} \tag{4.117}$$

where (a) is because $\widehat{f\psi} = \pi^* f\hat{\psi}$, and (b) is an application of the Leibniz rule. □

Theorem 4.28.
Using the local coordinates related to the sections $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$, the covariant derivatives reads:

$$(\nabla_X\psi)_{(\alpha)}(x) = X_x\psi_{(\alpha)} - \rho_*(\sigma^*_\alpha\omega_x(X))\psi_{(\alpha)}(x) \tag{4.118}$$

where $\rho_*\colon \mathcal{G} \to \mathrm{End}(V)$ is defined by

$$\rho_*(A) = \frac{d}{dt}\Big[\rho(e^{tA})\Big]_{t=0} \tag{4.119}$$

Proof. The problem reduces to the search of $\overline{X}$ because

$$(\nabla_X\psi)_{(\alpha)}(x) = \widehat{\nabla_X\psi}(\sigma_\alpha(x)) = \overline{X}_{\sigma_\alpha(x)}(\hat{\psi}).$$

We claim that $\overline{X}_{\sigma_\alpha(x)} = d\sigma_\alpha X_x - \omega(d\sigma_\alpha X_x)^*$. We have to check that $d\pi\overline{X} = X$ and $\omega(\overline{X}) = 0$. The latter comes easily from the fact that $\omega(A^*) = A$. For the first one, remark that s_α is a section, then $d(\pi \circ s_\alpha) = \mathrm{id}$, and $d\pi(ds_\alpha X_x) = X_x$, while

$$d\pi(A^*_\xi) = d\pi\frac{d}{dt}\Big[\xi \cdot e^{-tA}\Big]_{t=0} = \frac{d}{dt}\Big[\pi(\xi)\Big]_{t=0} = 0. \tag{4.120}$$

Since the horizontal lift is unique, we deduce

$$(\nabla_X\psi)_{(\alpha)}(x) = \big(d\sigma_\alpha X_x - \omega(d\sigma_\alpha X_x)^*\big)\hat{\psi}. \tag{4.121}$$

From the definition of a fundamental vector field,

$$\begin{aligned}\omega(d\sigma_\alpha X_x)^*_{\sigma_\alpha(x)}\hat{\psi} &= \frac{d}{dt}\Big[\hat{\psi}\big(\sigma_\alpha(x)\cdot e^{-t\omega(d\sigma_\alpha X_x)}\big)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\rho(e^{t\omega(d\sigma_\alpha X_x)})\hat{\psi}(\sigma_\alpha(x))\Big]_{t=0} \qquad \text{from (4.48)}\\ &= (d\rho)_e(\omega\circ d\sigma_\alpha)X_x(\hat{\psi}\circ\sigma_\alpha)(x)\\ &= \rho_*\big((\sigma_\alpha^*\omega)(X_x)\big)\psi_{(\alpha)}(x) \qquad \text{by (4.119)}\end{aligned} \tag{4.122}$$

□

We can express the covariant derivative by means of some maps $\theta_\alpha\colon \mathfrak{X}(M)\times M\to \mathrm{End}(V)$ given by

$$\nabla_X\gamma_{\alpha i} = \theta_\alpha(X)_i^{\ j}\gamma_{\alpha j}. \tag{4.123}$$

where the $\gamma_{\alpha i}$'s were given in equation (4.55). By the definition (4.111b),

$$\begin{aligned}(\nabla_X\psi)(x) &= (X\cdot s^i_\alpha)_x\gamma_{\alpha i}(x) + s^i_\alpha(x)(\nabla_X\gamma_{\alpha i})(x)\\ &= (X\cdot s^i_\alpha)_x\gamma_{\alpha i}(x) + s^i_\alpha(x)\theta_\alpha(X)_i^{\ j}\gamma_{\alpha j}(x).\end{aligned}$$

On the othre hand with the notations of equation (4.53), $\gamma_{\alpha j} = e_i$ and $X_x\gamma_{\alpha j} = 0$. Then equation (4.118) gives $\theta_\alpha(X) = \rho_*(\sigma_\alpha^*\omega_x(X))$, or

$$\theta_\alpha = \rho_*(\sigma_\alpha^*\omega_x). \tag{4.124}$$

4.11.1 Curvature on associated bundle

From the definition (4.49), it makes sense to define the curvature 2-form by

$$R(X,Y)\psi = \nabla_X\nabla_Y\psi - \nabla_Y\nabla_X\psi - \nabla_{[X,Y]}\psi.$$

It is also clear that $\psi_{(\alpha)}$ defines a section of the trivial vector bundle $F = M\times V$ by $x\to(x,\psi_{(\alpha)}(x))$, so one can define $\Omega_\alpha(X,Y)\colon \Gamma(M,E)\to\Gamma(M,E)$ by

$$\big(R(X,Y)\psi\big)_{(\alpha)} = \Omega_\alpha(X,Y)\psi_{(\alpha)}$$

and take back all the work around Bianchi because of the relation (4.118) which can be written as $(\nabla_X\psi)_{(\alpha)}(x) = X_x\psi_{(\alpha)} + \theta_\alpha(X)\psi_{(\alpha)}(x)$ and which is the same as in proposition 4.19.

4.11.2 Connection on frame bundle

4.11.2.1 General framework

The frame bundle was defined at page 150. Let $F\xrightarrow{p}M$ be a $\mathbb{K}$-vector bundle with some local trivialization $(\mathcal{U}_\alpha,\phi^E_\alpha)$ and the corresponding transition functions $g_{\alpha\beta}\colon\mathcal{U}_\alpha\cap\mathcal{U}_\beta\to GL(r,\mathbb{K})$. We consider $\pi\colon P\to M$, the frame bundle of F; it is a $GL(r,\mathbb{K})$-principal bundle. Let ∇ be a covariant derivative on F and θ_α, the associated matrices 1-form. The frame bundle is

$$P = \bigcup_{x\in M}(\text{frame of } F_x).$$

A connection is a $\mathcal{G}$-valued 1-form; in our case it is a map

$$\omega^\alpha_\xi\colon T_\xi\big(\pi^{-1}(\mathcal{U}_\alpha)\big)\to\mathfrak{gl}(r,\mathbb{K}).$$

We define our connection by, for $g\in GL(r,\mathbb{K})$, $x\in\mathcal{U}_\alpha$, $X_x\in T_xM$ and $A\in\mathfrak{gl}(r,\mathbb{K})$,

$$\omega^\alpha_{S_\alpha(x)\cdot g}\big(R_{g*}s_\alpha(x)_*X_x + A^*_{S_\alpha(x)\cdot g}\big) := A + \mathrm{Ad}(g^{-1})\theta_\alpha(X_x). \tag{4.125}$$

where $S_\alpha\colon\mathcal{U}_\alpha\to P$ is the section defined by the trivialization ϕ^P_α:

$$S_\alpha(x) = \{\overline{v}_\alpha = \phi^{E^{-1}}_\alpha(x,e_i)\}_{i=1,\ldots,r}.$$

Since $\theta_\alpha(X_x)\in\mathrm{End}(\mathbb{K}^r)\subset\mathfrak{gl}(r,\mathbb{K})$, the second term of (4.125) makes sense. This formula is a good definition of ω because of the following lemma:

Lemma 4.29.
If $\xi = S_\alpha(x) \cdot g$ and $\Sigma \in T_\xi P$, there exists a choice of $A \in \mathcal{G}$, and $X_x \in T_x M$ such that

$$\Sigma = R_{g*} s_\alpha(X)_* X_x + A^*_{S_\alpha(x)\cdot g}. \tag{4.126}$$

Proof. If $\xi \in P$ is a basis of E at y, there exists only one choice of $x \in M$ and $g \in G$ such that $\xi = S_\alpha(x) \cdot g$.

Let us consider a general path $c\colon \mathbb{R} \to P$ under the form $c(t) = s_\alpha(x(t)) \cdot g(t)$ where x and g are path in M and $GL(r, \mathbb{K})$. The frame $c(t)$ is the one of $F_{x(t)}$ obtained by the transformation $g(t)$ from $s_\alpha(x(t))$. It is a set of r vectors, and each of them can be written as a combination of the vectors of $s_\alpha(x(t))$, so we write

$$c^i(t) = s^j_\alpha(x(t)) g^i_j(t) \tag{4.127}$$

where $s^j_\alpha(x(t)) \in F_{x(t)}$ and $g^i_j(t) \in \mathbb{K}$. We compute $\Sigma = c'(0)$ by using the Leibnitz rule and we denote $x'(0) = X_x$, $x(0) = x$ and $g^i_j(0) = g^i_j$ (the matrix of g):

$$\begin{aligned}
\Sigma^i &= \frac{d}{dt}\Big[s^j_\alpha(x(t))\Big]_{t=0} g^i_j + s^j_\alpha(x) \frac{d}{dt}\Big[g^i_j(t)\Big]_{t=0} \\
&= (ds^j_\alpha)_x X_x g^i_j + {g^i_j}'(0) s^j_\alpha(x).
\end{aligned} \tag{4.128}$$

Going to more compact matrix form, it gives

$$\Sigma = (ds_\alpha)_x X_x \cdot g + s_\alpha(x) g'(0).$$

The second term, $s^j_\alpha(x) g'^i_j(0)$, is a general vector tangent to a fiber. So it can be written as a fundamental field A^*_ξ.

□

Lemma 4.30.
On $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$, the form fulfills $\omega^\alpha = \omega^\beta$.

Proof. Let $\gamma\colon \mathbb{R} \to M$ be a path whose derivative is X_x. Then

$$\begin{aligned}
(R_g)_* s_\alpha(x)_* X_x &= \frac{d}{dt}\Big[s_\alpha(\gamma_t) \cdot g\Big]_{t=0} = \frac{d}{dt}\Big[s_\beta(\gamma_t) g_{\alpha\beta}(\gamma_t) \cdot g\Big]_{t=0} \\
&= \frac{d}{dt}\Big[s_\alpha(\gamma_t) g_{\alpha\beta}(x) \cdot g\Big]_{t=0} + \frac{d}{dt}\Big[s_\beta(x) \cdot g_{\alpha\beta}(\gamma_t) \cdot g\Big]_{t=0}.
\end{aligned} \tag{4.129}$$

What is in the derivative of the first term is $R_{g_{\alpha\beta}(x)g}(s_\beta(\gamma_t))$. Taking the derivative, we find the expected $R_{g_{\alpha\beta}(x)g*} s_{\beta*} X_x$.

For the second term, we note $r := s_\beta(x) \cdot g_{\alpha\beta}(g) g$, and we have to compute the following, using equation (4.16),

$$\begin{aligned}
&\frac{d}{dt}\Big[r \cdot \operatorname{Ad}_{g^{-1}}(g^{-1}_{\alpha\beta}(x) g_{\alpha\beta}(\gamma_t))\Big]_{t=0} \\
&\qquad = \frac{d}{dt}\Big[r \cdot \exp t\big((d\,\mathbf{Ad}_{g^{-1}})_e (g^{-1}_{\alpha\beta}(x)(dg_{\alpha\beta})_x X_x)\big)\Big]_{t=0} \\
&\qquad = \frac{d}{dt}\Big[r \cdot \exp t\big(\operatorname{Ad}_{g^{-1}} g^{-1}_{\alpha\beta}(x) dg_{\alpha\beta}\Big]_{t=0} \\
&\qquad = \Big(\operatorname{Ad}_{g^{-1}} g^{-1}_{\alpha\beta}(x) dg_{\alpha\beta} X_x\Big)^*_r.
\end{aligned} \tag{4.130}$$

Using this, we can perform the computation:

$$\begin{aligned}
\omega^\beta_{S_\alpha(x)\cdot g}\big(R_{g*} s_\alpha(x)_* X_x + A^*_{S_\alpha(x)\cdot g}\big) &= \omega^\beta_{S_\beta(x)\cdot g_{\alpha\beta}(x) g}\Big(R_{g_{\alpha\beta}}(x) g_* s_\beta(x)_* X_x \\
&\qquad + (\operatorname{Ad}_{g^{-1}} g^{-1}_{\alpha\beta}(x) dg_{\alpha\beta} X_x)^*_r + A^*\Big) \\
&= \operatorname{Ad}_{(g_{\alpha\beta}(x)g)^{-1}} \theta_\beta(X_x) \\
&\quad + \operatorname{Ad}_{g^{-1}} g^{-1}_{\alpha\beta}(x) dg_{\alpha\beta}(X_x) + A \\
&= \operatorname{Ad}_{g^{-1}}\big((g^{-1}_{\alpha\beta}\theta_\beta g_{\alpha\beta} + g^{-1}_{\alpha\beta} dg_{\alpha\beta})(X_x)\big) + A \\
&= \omega^\alpha_{S_\alpha(x) g}\big(R_{g*} s_\alpha(x)_* X_x + A^A_{S_\alpha(x)\cdot g}\big).
\end{aligned} \tag{4.131}$$

□

Proposition 4.31.
The ω defined by formula (4.125) is a connection 1-form.

Proof. The first condition, $\omega(A^*_\xi) = A$, is immediate from the definition. The lemma 2.11 gives the second condition in the case $\Sigma = A^*_\xi$. It remains to be checked that $\omega(dR_g\Sigma) = \mathrm{Ad}(g^{-1})\omega(\Sigma)$ in the case $\Sigma = dR_h ds_\alpha X_x$. This is obtained using the fact that Ad is a homomorphism. □

4.11.2.2 Levi-Civita connection

Let (M,g) be a Riemannian manifold. We look at a connection 1-form $\alpha \in \Omega^1(\mathrm{SO}(M), so(\mathbb{R}^m))$ on $\mathrm{SO}(M)$, and we define a covariant derivative $\nabla^\alpha \colon \mathfrak{X}(M) \times T(M) \to T(M)$, where $T(M)$ is the tensor bundle on M by (cf. theorem (4.27))

$$\widehat{\nabla^\alpha_X s} = \overline{X}\hat{s}, \tag{4.132}$$

for any $s \in T(M)$. Our purpose now is to prove that an automatic property of this connection is $\nabla^\alpha g = 0$. The unique such connection which is torsion-free is the **Levi-Civita** one.

The metric g is a section of the tensor bundle $T^*M \otimes T^*M$. So we have, in order to find $\hat{g}$ and to use equation (4.132), to see $T^*M \otimes T^*M$ as an associated bundle. As done in 4.5.4, we see that

$$T^*M \otimes T^*M \simeq \mathrm{SO}(M) \times_\rho (V^* \otimes V^*),$$

with the following definitions:

- The isomorphism is given by $\psi[b, \alpha\otimes\beta](X\otimes Y) = \alpha(b^{-1}X)\beta(b^{-1}Y)$,
- $\rho(A)\alpha = \alpha \circ A$,
- $b \cdot A = b \circ A$.

Here, $V = \mathbb{R}^m$; $b \colon V \to T_xM$; $\alpha, \beta \in V^*$; $X, Y \in T_xM$ and $A \in \mathrm{SO}(m)$ is seen as $A \colon V \to V$.

The following shows that ψ is well defined:

$$\begin{aligned}\psi[b\cdot A, \rho(A^{-1})\alpha\otimes\beta](X\otimes Y) &= (\alpha\circ A)(A^{-1}\circ b^{-1}X)(\beta\circ A)(A^{-1}\circ b^{-1}Y)\\ &= \psi[b,\alpha\otimes\beta](X\otimes Y)\end{aligned} \tag{4.133}$$

Proposition 4.32.
The function $\hat{g}$ is given by

$$\hat{g}(b)(v\otimes w) = g_x(b(v)\otimes b(w)) = v\cdot w.$$

Proof. The second equality is just the fact that $b \colon (\mathbb{R}^m, \cdot) \to (T_xM, g_x)$ is isometric. On the other hand, if $\hat{g}(b) = \alpha\otimes\beta$, we have:

$$\begin{aligned}g_x(X\otimes Y) &= \psi[b,\alpha\otimes\beta](X\otimes Y) = \alpha(b^{-1}X)\beta(b^{-1}Y)\\ &= \alpha\otimes\beta(b^{-1}X\otimes b^{-1}Y) = \hat{g}(b)(b^{-1}X\otimes b^{-1}Y).\end{aligned} \tag{4.134}$$

Since b is bijective, X and Y can be written as bv and bw respectively for some $v, w \in V$, so that

$$g_x(bv\otimes bw) = \hat{g}(b)v\otimes w.$$

□

It is now easy to see that $\overline{X}\hat{g} = 0$. As $\hat{g}$ takes its values in $V^*\otimes V^*$, $\overline{X}\hat{g}$ belongs to this space and can be applied on $v\otimes w \in V\otimes V$. Let $\overline{X}(t)$ be a path in $\mathrm{SO}(M)$ which defines $\overline{X}$; if $\overline{X} \in T_b\,\mathrm{SO}(M)$, $\overline{X}(0) = b$. We have

$$\overline{X}\hat{g}(v\otimes w) = \frac{d}{dt}\hat{g}(\overline{X}(t))v\otimes w\Big|_{t=0} = \frac{d}{dt}v\cdot w\Big|_{t=0}, \tag{4.135}$$

which is obviously zero.

4.11.3 Holonomy

Let the principal bundle

$$\begin{tikzcd} G \arrow[r, squiggly] & P \arrow[d, "\pi"] \\ & M \end{tikzcd} \tag{4.136}$$

and ω a connection on G. Let $\gamma\colon [0,1] \to M$, a closed curve piecewise smooth; $\gamma(0) = \gamma(1) = x$. For each $p \in \pi^{-1}(x)$, there exists one and only one horizontal lift $\tilde{\gamma}\colon [0,1] \to P$ such that $\tilde{\gamma}(0) = p$. There exists of course an element $g \in G$ such that $\tilde{\gamma}(1) = p \cdot g$.

We define the following equivariance relation on P: we say that $p \sim q$ if and only if p and q can be joined by a piecewise smooth path. The **holonomy group** at the point p is

$$\mathrm{Hol}_p(\omega) = \{g \in G \text{ st } p \sim p \cdot g\}.$$

4.11.4 Connection and gauge transformation

Proposition 4.33.
If ω is a connection on a G-principal bundle and φ, a gauge transformation, the form $\beta = \varphi^*\omega$ is a connection 1-form too.

Proof. It is rather easy to see that $\varphi_* A^*_\xi = A^*_{\varphi(x)}$:

$$\varphi_* A^*_\xi = \frac{d}{dt}\Big[\varphi(\xi e^{-tA})\Big]_{t=0} = \frac{d}{dt}\Big[\varphi(\xi)e^{-tA}\Big]_{t=0} = A^*_{\varphi(x)}.$$

The same kind of reasoning leads to $\varphi_* R_{g*} = R_{g*}\varphi_*$. From here, it is easy to see that

$$(\varphi^*\omega)_\xi(A^*_\xi) = \omega_{\varphi(\xi)}(\varphi_* A^*_\xi) = A,$$

and

$$\big(R^*_g(\varphi^*\omega)_\xi\big)(\Sigma) = (R^*_g\omega)_{\varphi(\xi)}(\varphi_*\Omega) = \mathrm{Ad}(g^{-1})\big((\varphi^*\omega)_\xi(\Sigma)\big).$$

□

So, the "gauge transformed" of a connection is still a connection. It is hopeful in order to define gauge invariants objects (Lagrangian) from connections (electromagnetic fields).

4.11.4.1 Local description

Let $\pi\colon P \to M$ be a G-principal bundle given with some trivializations $\phi^P_\alpha\colon \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times G$ over $\mathcal{U}_\alpha \subset M$ and $s_\alpha\colon \mathcal{U}_\alpha \to \pi^{-1}(\mathcal{U}_\alpha)$, a section. In front of that, we consider an associated bundle $p\colon E = P \times_\rho V \to M$ with a trivialization $\phi^E_\alpha\colon E \to \mathcal{U}_\alpha \times V$. One can choose a section s_α compatible with the trivialization in the sense that $\phi^P_\alpha(s_\alpha(x) \cdot g) = (x, g)$; the same can be done with E by choosing $\phi^E_\alpha([s_\alpha(x), v]) = (x, v)$. All this given in figure 4.1.

A section $\psi\colon M \to E$ is described by a function $\psi_\alpha\colon \mathcal{U}_\alpha \to V$ defined by $\phi^E_\alpha(\psi(x)) = (x, \psi_\alpha(x))$. In the inverse sense, ψ is defined (on $\mathcal{U}_\alpha$) from ψ_α by $\psi(x) = [s_\alpha(x), \psi_\alpha(x)]$. In the same way, a gauge transformation $\varphi\colon P \to P$ is described by functions $\tilde{\varphi}_\alpha\colon \mathcal{U}_\alpha \to G$,

$$\varphi(s_\alpha(x)) = s_\alpha(x) \cdot \tilde{\varphi}_\alpha(x). \tag{4.137}$$

The function $\tilde{\varphi}_\alpha$ also fulfil

$$(\phi^P_\alpha \circ \varphi \circ \phi^{P^{-1}}_\alpha)(x, g) = (x, \tilde{\varphi}(x) \cdot g) \tag{4.138}$$

because

$$\begin{aligned}
(\phi^P_\alpha \circ \varphi \circ \phi^{P^{-1}}_\alpha)(x, g) &= (\phi^P_\alpha \circ \varphi)(s_\alpha(x) \cdot g) \\
&= \phi^P_\alpha(\varphi(s_\alpha(x)) \cdot g) \\
&= \phi^P_\alpha(s_\alpha(c) \cdot \tilde{\varphi}_\alpha(x)g) \\
&= (x, \tilde{\varphi}_\alpha(x)g).
\end{aligned} \tag{4.139}$$

We know that a connection on P is given by its 1-form ω. Moreover we have the following:

Proposition 4.34.
A connection on P is completely determined on $\pi^{-1}(\mathcal{U}_\alpha)$ from the data of the $\mathcal{G}$-valued 1-form $\sigma^*_\alpha\omega$ on $\mathcal{U}_\alpha$.

Proof. We consider a 1-form ω which fulfils the two conditions of page 163. Our purpose is to find back $\omega_\xi(\Sigma)$, $\forall \xi \in P, \Sigma \in T_\xi P$ from the data of $\sigma_\alpha^*\omega$ alone. For any ξ, there exists a g such that $\xi = \sigma_\alpha(x) \cdot g$. We have

$$\mathrm{Ad}_{g^{-1}}(\omega_{\sigma_\alpha(x)\Sigma}) = (R_g^*\omega)_{\sigma_\alpha(x)}(\Sigma) = \omega_{\sigma_\alpha(x)\cdot g}\big((dR_g)_{\sigma_\alpha(x)}\Sigma\big). \tag{4.140}$$

If we know $s_\alpha^*\omega$, then we know $\omega\big((ds_\alpha)_x v\big)$ for any $v \in T_x M$. So

$$\omega_{\sigma_\alpha(x)\cdot g}\big((dR_g)_{\sigma_\alpha(x)}\Sigma\big)$$

is given from $\sigma_\alpha^*\omega$ for every Σ of the form $\Sigma = (d\sigma_\alpha)_x v$. From the form (4.126) of a vector in $T_\xi P$, it just remains to express $\omega_{\sigma_\alpha(x)\cdot g}(A^*_{\sigma_\alpha(x)\cdot g})$ in terms of s_α^*. The definition of a connection makes that it is simply A. □

4.11.4.2 Covariant derivative

If we have a connection on P, we can define a covariant derivative on the associated bundle E by

$$(\nabla_X\psi)_{(\alpha)}(x) = X_x(\psi_\alpha) + \rho_*(s_\alpha^*\omega_x(X))\psi_{(\alpha)}(x),$$

the matricial 1-form being given by $\theta_\alpha = \rho_*\sigma_\alpha^*\omega$. The gauge transformation φ acts on the connection ω by defining $\omega^\varphi := \varphi^*\omega$.

Proposition 4.35.
If $\beta = \varphi^*\omega$, then

$$s_\alpha^*(\beta) = \mathrm{Ad}_{\tilde\varphi_\alpha(x)^{-1}}\, s_\alpha^*(\omega) + \tilde\varphi_\alpha(x)^{-1} d\tilde\varphi_\alpha.$$

Proof. Let $\gamma\colon \mathbb{R} \to M$ be a path such that $\gamma(0) = x$ and $\gamma'(0) = X_x$. We have to compute the following:

$$(s_\alpha^*\beta)(X_x) = (s_\alpha^*\varphi^*\omega)(X_x) = \omega_{(\varphi\circ s_\alpha)(x)}\big((\varphi\circ s_\alpha)_* X_x\big). \tag{4.141}$$

What lies in the derivative is:

$$\begin{aligned}
(\varphi\circ s_\alpha)_*(X_x) &= \frac{d}{dt}\Big[(\varphi\circ s_\alpha\circ\gamma)(t)\Big]_{t=0}\\
&= \frac{d}{dt}\Big[s_\alpha(\gamma(t))\cdot\tilde\varphi_\alpha(\gamma(t))\Big]_{t=0}\\
&= \frac{d}{dt}\Big[s_\alpha(\gamma(t))\cdot\tilde\varphi_\alpha(\gamma(0))\Big]_{t=0} + \frac{d}{dt}\Big[s_\alpha(\gamma(0))\cdot\tilde\varphi_\alpha(\gamma(t))\Big]_{t=0}\\
&= R_{\tilde\varphi_\alpha(x)*}s_{\alpha*}X_x + \frac{d}{dt}\Big[s_\alpha(x)\cdot\tilde\varphi_\alpha(x)e^{t\tilde\varphi_\alpha(x)^{-1}(d\tilde\varphi_\alpha)_x\gamma'(0)}\Big]_{t=0}.
\end{aligned} \tag{4.142}$$

A justification of the remplacement $\tilde\varphi_\alpha(\gamma(t)) \to \tilde\varphi_\alpha(x)e^{t\tilde\varphi_\alpha(x)^{-1}(d\tilde\varphi_\alpha)_x\gamma'(0)}$ is given in the corresponding proof at page 245. If we put this expression into equation (4.141), the first term becomes

$$\begin{aligned}
\omega_{(\varphi\circ s_\alpha)(x)}\big(R_{\tilde\varphi_\alpha(x)*}s_{\alpha*}X_x\big) &= (R^*_{\tilde\varphi_\alpha(x)}\omega)_{s_\alpha(x)}(s_{\alpha*}X_x)\\
&= \mathrm{Ad}_{\tilde\varphi_\alpha(x)^{-1}}\big(\omega_{s_\alpha(x)}(s_{\alpha*}X_x)\big)\\
&= \mathrm{Ad}_{\tilde\varphi_\alpha(x)^{-1}}(s_\alpha^*\omega)(X_x).
\end{aligned}$$

The second term is the case of a connection applied to a fundamental vector field.

□

4.12 Product of principal bundle

In this section, we build a $G_1 \times G_2$-principal bundle from the data of a G_1 and a G_2-principal bundle. The physical motivation is clear: as far as electromagnetism is concerned, particles are sections of $U(1)$-principal bundle while the relativistic invariance must be expressed by means of a $\mathrm{SL}(2,\mathbb{C})$-associated bundle. So the physical fields must be sections of something as the product of the two bundles. See subsection 7.5.

4.12.1 Putting together principal bundle

Let us consider two principal bundle over the same base space

$$G_1 \rightsquigarrow P_1 \xrightarrow{p_1} M,$$

and

$$G_2 \rightsquigarrow P_2 \xrightarrow{p_2} M.$$

First we define the set

$$P_1 \circ P_2 = \{(\xi_1, \xi_2) \in P_1 \times P_2 \text{ st } p_1(\xi_1) = p_2(\xi_2)\} \tag{4.143}$$

which will be the total space of our new bundle. The projection $p\colon P_1 \circ P_2 \to M$ is naturally defined by

$$p(\xi_1, \xi_2) = p_1(\xi_1) = p_2(\xi_2),$$

while the right action of $G_1 \times G_2$ on $P_1 \circ P_2$ is given by

$$(\xi_1, \xi_2) \cdot (g_1, g_2) = (\xi_1 \cdot g_1, \xi_2 \cdot g_2)$$

With all these definitions,

$$\begin{tikzcd} G_1 \times G_2 \arrow[r, rightsquigarrow] & P_1 \circ P_2 \arrow[d, "p"] \\ & M \end{tikzcd}$$

is a $G_1 \times G_2$-principal bundle over M. We define the natural projections

$$\pi_i\colon P_1 \times P_2 \to P_i \ (\xi_1, \xi_2) \quad \mapsto \xi_i, \tag{4.144}$$

and if e_i denotes the identity element of G_i, we can identify G_1 to $G_1 \times \{e_2\}$ and G_2 to $G_2 \times \{e_1\}$; in the same way, $\mathcal{G}_1 = \mathcal{G}_1 \times \{0\} \subset \mathcal{G}_1 \times \mathcal{G}_2$. So we get the following principal bundles:

$$G_2 \rightsquigarrow P_1 \circ P_2 \xrightarrow{\pi_1} P_1$$

$$G_1 \rightsquigarrow P_1 \circ P_2 \xrightarrow{\pi_2} P_2.$$

It is clear that the following diagram commutes:

$$\begin{tikzcd} P_1 \arrow[dr, "P_1"'] & P_1 \circ P_2 \arrow[l, "\pi_1"'] \arrow[r, "\pi_2"] \arrow[d, "p"] & P_2 \arrow[dl, "p_2"] \\ & M & \end{tikzcd}$$

4.12.2 Connections

Let ω_i be a connection on the bundle $p_i\colon P_i \to M$. Using the identifications, $\pi_1^*\omega_1$ is a connection on $\pi_2\colon P_1 \circ P_2 \to P_2$ (the same is true for $1 \leftrightarrow 2$), and $\pi_1^*\omega_1 \oplus \pi_2^*\omega_2$ is a connection on $p\colon P_1 \circ P_2 \to M$. Let us prove the first claim.

Let $A \in \mathcal{G}_1$. We first have to prove that $\pi_1^*\omega_1(A^*) = A$. For this, remark that $A = (A, 0) \in \mathcal{G}_1 \oplus \mathcal{G}_2$ and

$$A^*_\xi = \frac{d}{dt}\Big[\xi \cdot e^{-t(A,0)}\Big]_{t=0} = \frac{d}{dt}\Big[(\xi_1, \xi_2) \cdot (e^{-tA}, e_2)\Big]_{t=0} = \frac{d}{dt}\Big[(\xi_1 \cdot e^{-tA}, \xi_2)\Big]_{t=0}, \tag{4.145}$$

so $d\pi_1 A^* = \frac{d}{dt}\Big[\pi_1(\ldots)\Big]_{t=0} = \omega_1(A^*) = A$. Let now $\Sigma \in T_{(\xi_1,\xi_2)}(P_1 \circ P_2)$ be given by the path $(\xi_1(t), \xi_2(t))$. In this case we have

$$\begin{aligned}
\big(R^*_{(g,e_2)\pi_1^*\omega_1}\big)_{(\xi_1,\xi_2)}\Sigma &= (\pi_1^*\omega_1)(dR_{(g,e_2)}\Sigma) \\
&= \omega_1\big(\frac{d}{dt}\Big[\xi_1(t) \cdot g\Big]_{t=0}\big) \\
&= \omega_1\big(dR_g \frac{d}{dt}\Big[\xi_1(t)\Big]_{t=0}\big) \\
&= \operatorname{Ad}(g^{-1})\pi_1^*\omega_1\big(\frac{d}{dt}\Big[(\xi_1(t), \xi_2(t))\Big]_{t=0}\big) \\
&= \operatorname{Ad}(g^{-1})\pi_1^*\omega_1\Sigma.
\end{aligned} \tag{4.146}$$

4.12.3 Representations

Let V be a vector space and $\rho_i\colon G_i \to GL(V)$ be some representations such that

$$[\rho_1(g_1), \rho_2(g_2)] = 0 \tag{4.147}$$

for all $g_1 \in G_1$ and $g_2 \in G_2$ (in the sense of commutators of matrices). In this case, one can define the representation $\rho_1 \times \rho_2\colon G_1 \times G_2 \to GL(V)$ by

$$(\rho_1 \times \rho_2)(g_1, g_2) = \rho_1(g_1) \circ \rho_2(g_2) = \rho_2(g_2) \circ \rho_1(g_1). \tag{4.148}$$

The relation (4.147) is needed in order for $\rho_1 \times \rho_2$ to be a representation, as one can check by writing down explicitly the requirement

$$(\rho_1 \times \rho_2)\big((g_1, g_2)(g_1', g_2')\big) = (\rho_1 \times \rho_2)(g_1g_1', g_2g_2')$$

4.13 Hodge decomposition theorem and harmonic forms

Among other sources for Hodge decomposition and harmonic forms, we have [33–35]. Some parts of the wikipedia article Hodge_dual are interesting.

Let E be an oriented Euclidian space of dimension $m = 2n$. We define the operation $*$ by

$$\begin{aligned} *\colon \bigwedge E &\to \bigwedge E \\ e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k} &\mapsto e_{i_{k+1}} \wedge \cdots \wedge e_{i_m} \end{aligned} \tag{4.149}$$

when $\{e_i\}$ is an oriented basis of E and $\{i_k\}$ is an even permutation of $\{1, 2, \cdots, m\}$. If it is impossible to build an even permutation, then we add a minus sign. We have $**\omega = (-1)^{p(m-p)}\omega$ belongs to $\omega \in \bigwedge^p E$.

Example 4.36.
If we consider the space $\mathbb{R}^4$ with the coordinates (x, y, z, t),

$$*(dx \wedge dz \wedge dt) = dy \tag{4.150}$$

because (x, z, t, y) is an even permutation of (x, y, z, t). Now, $*dy = dz \wedge dx \wedge dt = -(dx \wedge dz \wedge dt)$ because (y, z, x, t) is an even permutation of (x, y, z, t).

More generally, if we have a differential p-form ω on a m dimensional space, we have

$$*\left(e_{\sigma(1)} \wedge e_{\sigma(2)} \wedge \ldots \wedge e_{\sigma(p)}\right) = e_{\sigma(p+1)} \wedge \ldots \wedge e_{\sigma(m)} \tag{4.151}$$

In order to compute $*(e_{\sigma(p+1)} \wedge \ldots \wedge e_{\sigma(m)})$, we need a permutation of $\big(\sigma(1), \ldots, \sigma(m)\big)$ which *begins* by $\sigma(p+1)\ldots\sigma(m)$. This reduces to permute the $m-p$ elements $\sigma(p+1), \ldots, \sigma(m)$ with the p first elements. Thus we have

$$**\omega = (-1)^{p(m-p)}\omega. \tag{4.152}$$

♢

Let V be a compact, oriented manifold. Each of the spaces of sections $C^\infty\big(V, \bigwedge_{\mathbb{C}}^k(T^*V)\big)$ is endowed with a 2-form

$$\langle \omega_1, \omega_2 \rangle = \int_V \omega_1 \wedge *\omega_2. \tag{4.153}$$

Lemma 4.37.
*The **codifferential** δ defined by*

$$\begin{aligned} \delta\colon \bigwedge_{\mathbb{C}}^k(T^*V) &\to \bigwedge_{\mathbb{C}}^{k-1}(T^*V) \\ \beta &\mapsto (-1)^{mk+m+1} * d * \beta \end{aligned} \tag{4.154}$$

is a formal adjoint of d for the product (4.153).

Proof. If $\beta \in \bigwedge_{\mathbb{C}}^k(T^*V)$ and $\alpha \in \bigwedge_{\mathbb{C}}^{k-1}(T^*V)$, we have

$$\begin{aligned} *\delta\beta &= (-1)^{mk+m+1} **\big(d * \beta\big) \\ &= (-1)^{mk+m+1}(-1)^{(m-k+1)(m-m+k-1)} d * \beta \\ &= (-1)^k d * \beta. \end{aligned} \tag{4.155}$$

Using that formula we find

$$\begin{aligned}
\langle d\alpha, \beta\rangle - \langle \alpha, \delta\beta\rangle &= \int_V d\alpha \wedge *\beta - \alpha \wedge *\delta\beta \\
&= \int_V d\alpha \wedge *\beta - (-1)^k \alpha \wedge d * \beta \\
&= \int_V d\alpha \wedge *\beta + (-1)^{k+1} \alpha \wedge d * \beta \\
&= \int_V d(\alpha \wedge *\beta) \\
&= \int_{\partial V} \alpha \wedge *\beta \\
&= 0.
\end{aligned} \tag{4.156}$$

This proves that $\langle d\alpha, \beta\rangle = \langle \alpha, \delta\beta\rangle$.

Problem and misunderstanding **20.**
I do not understand why the integral in the boundary is zero.

□

Now we define the **Laplace-Beltrami operator** by

$$\Delta = \delta d + d\delta \tag{4.157}$$

and the space of **harmonic forms**

$$H^k = \{\omega \in \Omega^k \text{ st } \Delta\omega = 0\}. \tag{4.158}$$

Lemma 4.38.
If M is a closed manifold, a k-form is harmonic if and only if $d\omega = \delta\omega = 0$.

Proof. No proof. □

Theorem 4.39 (Hodge decomposition theorem)**.**
For every integer $0 \leqslant k \leqslant m$, the space H^p is finite dimensional and $\Omega^k(M)$ has the orthogonal decomposition

$$\Omega^k(M) = H^k \oplus \Delta\big(\Omega^k(M)\big), \tag{4.159}$$

i.e. the space splits into the kernel of Δ and its image.

Theorem 4.40.
Let M be a compact orientable manifold of dimension m. Any exterior differential k-form can be written as a unique sum of an exact form, a coexact form and an harmonic form :

$$\omega = d\alpha + \delta\beta + \gamma. \tag{4.160}$$

with $\omega \in \Omega^k(M)$, $\alpha \in \Omega^{k-1}(M)$, $\beta \in \Omega^{k+1}(M)$ and $\gamma \in \Omega^k(M)$ harmonic.

The operator Δ commutes with the differential d and we have $d\Delta = \Delta d = d\delta d$ since

$$d\Delta\omega = dd\delta\omega + d\delta d\omega = d\delta d\omega, \tag{4.161}$$

because $d^2 = 0$, while

$$\Delta d\omega = d\delta d\omega + \Delta dd\omega = d\delta d\omega. \tag{4.162}$$

Lemma 4.41.
On a close manifold, $\Delta\omega = 0$ if and only if $\delta\omega = d\omega = 0$.

In the case of a closed manifold, a form is harmonic if and only if is belongs to the kernel of $d + \delta$. Moreover, a form in the image of $d + \delta$ is orthogonal to the harmonic forms :

$$\langle d\alpha^{k-1} + \delta\beta^{k+1}, \gamma\rangle = 0 \tag{4.163}$$

whenever γ is harmonic on a closed manifold.

Chapter 5

From Clifford algebras to Dirac operator

Bibliography for Clifford algebras, spin group and related topics are [36–40]. More agebraic point of view can be found in [41, 42]. More details about "square rooting" second order differential operators are in [43]. For physical concerns, the reader should refer to [44–46].

5.1 Invitation : Clifford algebra in quantum field theory

5.1.1 Schrödinger, Klein-Gordon and Dirac

The origin of the Klein-Gordon equation is almost the same as the one of the Schrödinger: one replace physical functions by operators. For a free particle, the correspondence are

$$\begin{aligned} &\text{energy} && E \to i\hbar\frac{\partial}{\partial t}, \\ &\text{momentum} && \mathbf{p} \to -i\hbar\nabla. \end{aligned}$$

The Schrödinger equation (which is the non relativistic quantum wave equation) comes from replacement in the non non relativistic expression of the Hamiltonian

$$E = \frac{\mathbf{p}^2}{2m} \longrightarrow \left(\partial_t - \frac{i\hbar}{2m}\nabla^2\right)\psi = 0,$$

while the Klein-Gordon one (which is the relativistic quantum wave equation) comes from the relativistic corresponding equation:

$$E^2 = \mathbf{p}^2c^2 + m^2c^4 \longrightarrow \left(\partial^\mu\partial_\mu + (\frac{mc}{\hbar})^2\right)\psi = 0.$$

This is a second order differential equation; there are however no "law of nature" which forbid a first order equation. We try

$$i\hbar\frac{\partial\psi}{\partial t} = \left(\frac{\hbar c}{i}\alpha^k\partial_k + \beta mc^2\right)\psi \equiv \hat{H}\psi.$$

There are some physical constraints on the coefficients α^k and β. We will study one of them: we want the components of ψ to satisfy the Klein-Gordon equation, so that the plane waves fulfill the fundamental relation $E^2 = p^2c^2 + m^2c^4$.

In order to see the implications of this constraint on the coefficients, we apply two times the operator $\hat{H}$, and we compare the result with the Klein-Gordon equation. We find:

$$\alpha^i\alpha^j + \alpha^j\alpha^i = 2\delta^{ij}\mathbb{1}, \tag{5.1a}$$

$$\alpha^i\beta + \beta\alpha^i = 0, \tag{5.1b}$$

$$(\alpha^i)^2 = \beta^2 = \mathbb{1}. \tag{5.1c}$$

If we define $\gamma^0 = \beta$ and $\gamma^i = \beta\alpha^i$, we find that the matrices γ^μ have to give a representation of the Clifford algebra[1]:

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{1}. \tag{5.2}$$

[1]Don't be afraid with the extra minus sign: the quantum field theory is most written with the metric $(+,-,-,-)$ instead of $(-,+,+,+)$.

The Dirac equation reads

$$\left(-i\gamma^\mu\partial_\mu + \frac{mc}{\hbar}\right)\psi = 0.$$

If we want to perform some computation with the quantum field theory, we need an explicit form for the γ's; that's the reason why we study representations of the Clifford algebra. The **Dirac operator** $\mathcal{D}$ is the operator which lies in the Dirac equation:

$$\mathcal{D} = \sum_{\mu=0}^{3}\gamma^\mu\frac{\partial}{\partial x^\mu}. \tag{5.3}$$

5.1.2 Lorentz algebra

There is an other physical reason (which is in fact the same, but differently presented) justifying the study of the Clifford algebra. The quantum field theory need representation of the Lorentz algebra[2]

$$[J^{\mu\nu}, J^{\rho\sigma}] = i(\eta^{\nu\rho}J^{\mu\sigma} - \eta^{\mu\rho}J^{\nu\sigma} - \eta^{\nu\sigma}J^{\mu\rho} + \eta^{\mu\sigma}J^{\nu\rho}).$$

A proof of these relations is given in lemma 5.1. Dirac had a trick to find such J matrices from a representation of the Clifford algebra. If we have $n\times n$ matrices γ_μ such that

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\mathbb{1}_{n\times n},$$

a n-dimensional representation of the Lorenz algebra is obtained by

$$S^{\mu\nu} = \frac{i}{4}\left[\gamma^\mu, \gamma^\nu\right].$$

Lemma 5.1.
The matrices of $\mathfrak{so}(p,q)$ satisfy the definition relation

$$M^t\eta + \eta M = 0, \tag{5.4}$$

and if M^{ab} is the "rotation" in the place of directions a and b (i.e. a trigonometric or an hyperbolic rotation following that a and b are of the same type or not), then the action on $\mathbb{R}^{(p,q)}$ is given by $(x')^\mu = (M^{ab})^\mu_\nu x^\nu$ with

$$(M^{ab})^\mu_\nu = \eta^{a\mu}\delta^b_\nu - \eta^{b\mu}\delta^a_\nu. \tag{5.5}$$

The commutation relations are given by

$$[M^{ab}, M^{cd}] = -\eta^{ac}M^{bd} + \eta^{ad}M^{bc} + \eta^{bc}M^{ad} - \eta^{bd}M^{ac}. \tag{5.6}$$

Notice that $M^{ab} = -M^{ba}$.

See section 12.5 of [47]. By a simple redefinition $J = iM$, one obtains

$$[J,J] = i\eta J \tag{5.7}$$

instead of $[M,M] = \eta M$, and the matrices J are Hermitian. Here η is the matrix $\eta = diag(\underbrace{+,\ldots,+}_{p\text{ times}},\underbrace{-,\ldots,-}_{q\text{ times}})$.
As convention, we say that a direction corresponding to a *positive* entry in the metric is a *time* direction, while the spatial directions are negative. That corresponds to the convention of page **??** to say that a *time-like* vector has positive norm.

5.2 Clifford algebra

5.2.1 Definition and universal problem

Definition 5.2.
Let V be a (finite dimensional) vector space and q, a bilinear quadratic form over V. The ***Clifford algebra*** $\mathrm{Cl}(V,q)$ *is the unital associative algebra generated by V subject to the relation*

$$v\cdot v = q(v) \tag{5.8}$$

for all v in $\mathrm{Cl}(V,q)$. Here the dot denotes the algebra product and $q(v)$ means $q(v,v)$.

[2]When one think to real infinitesimal rotation matrices, the presence of i seems not natural, but one redefines $J \to iJ$ for formalism reasons.

Theorem 5.4 proves unicity of such an algebra, so that it makes sense.

Remark 5.3.
The relation (5.8) *is no more a restriction for the elements in* $\mathrm{Cl}(V,q)$ *than a restriction on the choice of the algebra product.*

Theorem 5.4.
Let E be an unital associative algebra and $j\colon V \to E$ a linear map such that

$$j(v) \cdot j(v) = q(v)1. \tag{5.9}$$

Then we have an unique extension of j to a homomorphism $\tilde{j}\colon \mathrm{Cl}(V,q) \to E$. Moreover, $\mathrm{Cl}(V,q)$ is the unique associative algebra which have this property for all such E.

$$\begin{tikzcd} \mathrm{Cl}(V,q) \arrow[d, hook, "i"'] \arrow[dr, "\tilde{j}"] & \\ V \arrow[r, "j"'] & D \end{tikzcd}$$

This theorem can be seen as a definition of $\mathrm{Cl}(V,q)$.

Proof. The proof shall belongs two parts: the first one will show how to extend j and why it is unique, and the second one will prove the unicity of $\mathrm{Cl}(V,q)$.

We begin by define the extension of j. First note that any linear map $f\colon V \to E$ can be extended to an algebra homomorphism $\overline{f}\colon T(V) \to E$ in only one way. Indeed, the homomorphism condition require that $\overline{f}(v \otimes w) = f(v) \cdot f(w)$. The whole map $\overline{f}$ is then well defined by the data of f alone.

As far as the map j is concerned, we have the relation (5.9) which says that $\overline{j}(\mathcal{I}) = 0$. Indeed,

$$\overline{j}(v \otimes v - q(v) \cdot (1)) = \overline{j}(v) \cdot \overline{j}(v) - q(v)\overline{j}(1) = j(v) \cdot j(v) - q(v)1 = 0. \tag{5.10}$$

Thus $\overline{j}\colon T(V) \to E$ is a class map for $\mathcal{I}$, and we can descent $\overline{j}$ from $T(V)$ to $\mathrm{Cl}(V,q)$, We define $\tilde{j}\colon \mathrm{Cl}(V,q) \to E$ by

$$\tilde{j}[x] = \overline{j}(x) \tag{5.11}$$

where $[x]$ is the class of x. That's for the existence part.

The unicity is clear: $f_1 = f_2$ on V implies that $\overline{f_1} = \overline{f_2}$ on $T(V)$. Thus $\tilde{f}_1 = \tilde{f}_2$ on $\mathrm{Cl}(V,q)$.

We turn now our attention to the unicity of $\mathrm{Cl}(C,q)$. Let D be an unital associative algebra such that

(i) $V \subset D$,

(ii) For any unital associative algebra E and for any $f\colon D \to E$ such that $f(v) \cdot f(v) = -q(v)1$, there exists only one homomorphic map $\tilde{f}\colon D \to E$ which extend f.

We should find a homomorphic map $\tilde{k}\colon D \to \mathrm{Cl}(V,q)$. Let i be the canonical injection $i\colon V \to D$. Clearly, we have a homomorphism $V \to i(V)$. Now, as a space E, we can take $\mathrm{Cl}(V,q)$; i can be seen as a linear map $i\colon V \to \mathrm{Cl}(V,q)$ such that $i(v) \cdot i(v) = q(v)1$. The assumptions say that i can be extended (in only one way) to a homomorphic map $\tilde{i}\colon D \to \mathrm{Cl}(V,q)$.

The Clifford algebra is thus unique up to a homomorphism.

□

What we proved is the following: if for any E and for any $j\colon V \to E$ such that $j(v) \cdot j(v) = q(v)1$, there exist an unique $\tilde{j}\colon D \to E$ which extend j, then $D = \mathrm{Cl}(V,q)$ up to a homomorphism. One ays that $\mathrm{Cl}(V,q)$ solve an **universal problem**.

An explicit construction of $\mathrm{Cl}(V,q)$ can be achieved in the following way. We consider the tensor algebra $T(V) = \bigoplus_{n \geqslant 0} (\otimes^n V) = \mathbb{C} \oplus V \oplus (V \otimes V) \oplus \ldots$ over V the two-sided ideal $\mathcal{I}$ generated by elements of the form $v \otimes v - q(v)1$. The **Clifford algebra** for (V,q) is given by

$$\mathrm{Cl}(p,q) := T(V)/\mathcal{I} \tag{5.12}$$

in which product of $\mathrm{Cl}(V,q)$ is naturally defined by $[a] \otimes [b] = [a \otimes b]$ if $[a]$ is the class of $a \in T(V)$.

Let us now fix some notations more adapted to what we want to do. Let $V = \mathbb{R}^{p,q}$ the vector space $\mathbb{R}^{p+q}$ endowed with a diagonal metric which contains p plus sign and q minus signs. For v, $w \in V$, the inner product with respect to the metric η of v by w will be denoted by $\eta(v,w)$. The norm on V will be defined by

$\|v\|^2 = -\eta(v, v)$. It is neither positive defined, nor negative defined. The explanation of the minus sign will come soon. The Clifford algebra is the quotient $\mathrm{Cl}(p, q) := T(V)/\mathcal{I}$ of the tensor algebra by the two-sided ideal $\mathcal{I}$ generated by elements of the form

$$(v \otimes w) \oplus (w \otimes v) \oplus 2\eta(v, w)1$$

for v, w in V. Depending on the context, we will often use the notations $\mathrm{Cl}(\eta)$ or $\mathrm{Cl}(V)$ or $\mathrm{Cl}(p, q)$. The algebra product is $[x] \cdot [y] = [x \otimes y]$, x, $y \in T(V)$. As long as $z \in V \subset \mathrm{Cl}(p, q)$, the expression $\eta(z, z)$ is meaningful. The definition of Cl is such that $z \cdot z = -\eta(z, z)$. This leads to the somewhat surprising formula $z^2 = \|z\|^2 = -\eta(z, z)$.

5.2.2 First representation

Let (V, g) be a metric vector space and $\mathrm{Cl}(V, g)$ its Clifford algebra. For each $v \in V$, we define the two following elements of $\mathrm{End}_{\mathbb{R}}(\bigwedge V)$:

$$\epsilon(v)\big(u_1 \wedge \cdots \wedge u_k\big) = v \wedge u_1 \wedge \cdots \wedge u_k \tag{5.13a}$$

$$\iota(v)\big(u_1 \wedge \cdots \wedge u_k\big) = \sum_{j=1}^{k} (-1)^{j-1} g(u, u_j) u_1 \wedge \cdots \wedge \hat{u}_j \wedge \cdots \wedge u_j. \tag{5.13b}$$

One has $\epsilon(v)^2 = 0$ and $\iota(v)^2 = 0$ because $v \wedge v = 0$. In order to understand the latter, we wonder what are the terms with $g(v, u_i)g(v, u_j)$ are in

$$\iota(v)^2\big(u_1 \wedge \cdots \wedge u_k\big) = \sum_{l=1}^{k} (-1)^{j-1} g(v, u_j) \sum_{l=1}^{k-1} (-1)^{l-1} g(v, u_l) u_1 \wedge \hat{u}_l \wedge \hat{u}_j \wedge \cdots \wedge u_k.$$

Let's suppose $i < j$. The first term comes when the first $\iota(v)$ acts on u_j, its sign is given by $(-1)^{j-1}(-1)^{i-1}$. The second term has the same $(-1)^{i-1}$, but in this term, u_j is on the position $j - 1$ because u_i has disappeared.

Now we use $c(v) = \epsilon(v) + \iota(v)$ which fulfils for all u, $v \in V$:

$$\begin{aligned} c(v)^2 &= g(v, v,)1 \\ c(u)v(v) + c(v)c(u) &= 2g(u, v)1. \end{aligned}$$

Therefore c can be extended to a representation $c\colon \mathrm{Cl}(V, g) \to \mathrm{End}(\bigwedge V)$. If $\{e_0, \cdots e_n\}$ is an orthonormal basis of V (i.e. $g(e_i, e_j) = \eta_{ij}$); in this case the $c(e_j)$ are anticommuting and a basis of $\mathrm{Cl}(V, g)$ is given by

$$\{c(e_{k_1}) \cdots c(e_{k_r}) \text{ st } 1 \leqslant k_1 < \cdots < k_r \leqslant n\}. \tag{5.14}$$

5.2.3 Some consequences of the universal property

The map $-\mathrm{id}|_V$ extends to $\alpha \in \mathrm{Aut}\big(\mathrm{Cl}(V)\big)$,

$$\alpha(v_1 \cdots v_r) = (-1)^r v_1 \cdots v_r$$

($v_i \in V$) and provides a graduation

$$\mathrm{Cl}(V) = \mathrm{Cl}^0(V) \oplus \mathrm{Cl}^1(V).$$

The map $\tau\colon \mathrm{Cl}(V) \to \mathrm{Cl}(V)$ extends $\mathrm{id}|_V$ to an anti-homomorphism:

$$\tau(v_1 \cdots v_r) = v_r \cdots v_1. \tag{5.15}$$

The **complexification** of $\mathrm{Cl}(V, g)$ is

$$\mathrm{Cl}^{\mathbb{C}}(V, g) := \mathrm{Cl}(V, g) \otimes_{\mathbb{R}} \mathbb{C} \simeq \mathrm{Cl}(V^{\mathbb{C}}, g^{\mathbb{C}}),$$

the isomorphism being a $\mathbb{C}$-algebra isomorphism. The $\mathbb{R}$-linear operator $v \mapsto \overline{v}$ in $V^{\mathbb{C}}$ of complex conjugation extends to a $\mathbb{R}$-linear automorphism $a \mapsto \overline{a}$. We define the **adjoint** by

$$a^* = \tau(\overline{a}) \tag{5.16}$$

5.2.4 Trace

Theorem 5.5.
There exists one an only one trace $\operatorname{Tr}\colon \operatorname{Cl}^{\mathbb{C}}(V) \to \mathbb{C}$ *such that*

(i) $\operatorname{Tr}(1) = 1$,

(ii) $\operatorname{Tr}(a) = 0$ *when* a *is odd.*

Proof. Let $\{e_1, \cdots, e_n\}$ be an orthonormal basis of (V, g) and $a \in \operatorname{Cl}^{\mathbb{C}}(V)$. When one decomposes a into the basis of e_i, one finds a lot of terms of each order. Since Tr is a trace, when the k_i are all different,

$$\operatorname{Tr}(e_{k_1} \cdots e_{k_{2r}}) = \operatorname{Tr}(-e_{k_2} \cdots e_{k_{2r}} e_{k_1} = \operatorname{Tr}(-e_{k_1} \cdots e_{k_{2r}})$$

So the trace of any even element is zero. We decompose a into

$$a = \sum_K a_K \prod_{i \in K} e_i$$

where the sum is taken on the subsets of $\{1, \ldots, n\}$. A trace which fulfils the conditions must vanishes on even (but non zero) elements as well as on odd elements, so the only possible form is

$$\operatorname{Tr} a = a_{\varnothing}.$$

Notice that in order to get this precise form, we used $\operatorname{Tr}(1) = 1$ and linearity. This proves unicity and existence. Now we have to prove that this is a good definition in the sense that an other choice of basis gives the same result. So we take a new orthonormal basis

$$e'_j = \sum_{k=1}^{n} H_{jk} e_k$$

with $H^t H = \mathbb{1}_{n \times n}$. Now we have

$$a = \sum_K a_K \prod_{i \in K} e_i = \sum_K a'_K \prod_{i \in K} e'_i,$$

and we will prove that $a_{\varnothing} = a'_{\varnothing}$. Let's compute a lot:

$$\begin{aligned}
e'_i e'_j &= \sum_k \sum_l H_{ik} H_{jl} e_k e_l \\
&= \sum_{k=l} H_{ik} H_{jl} e_k e_l + \sum_{k \neq l} H_{ik} H_{jl} e_k e_l \\
&= \sum_k H_{ik} H_{jk} 1 + \sum_{k \neq l} H_{ik} H_{jl} e_k e_l \\
&= (HH^t)_{ij} 1 + \sum_{k \neq l} H_{ik} H_{jl} e_k e_l.
\end{aligned}$$

The sense of this formula is that when $i \neq j$, the product $e'_i e'_j$ has no term of order zero. In other terms, as long as we only have terms of order zero, one and two, a change $e \to e'$ does not change the term of order zero. We are now going to an induction proof: we want to prove that $e'_{j_1} \ldots e'_{j_{2r}} e'_l e'_k$ has no scalar term assuming that no even combination has scalar terms up to $2(r-1)$. It reads

$$\sum_{K \text{ even}} a_K \prod_{i \in K} e_i e'_l e'_k,$$

therefore we just have to look at terms of the form

$$e_{j_1} \ldots e_{j_{2r}} \Big((HH)^t_{kl} 1 - \sum_{i \neq j} C^{ij}_{kl} e_i e_j \Big)$$

where the e_{j_l} are all different. The first term cannot produce a scalar term. In order to find a scalar term in $e'_{j_1} \ldots e'_{j_{2r}} e_k e_l$, we begin to look at terms whose decomposition of $e'_{j_1} \ldots e'_{j_{2r}}$ ends by $e_l e_k$, i.e.

$$H_{j_{2r-2} l} H_{j_{2r-1} k} e'_{j_1} \ldots e'_{2r-3} e_l e_k e_k e_l.$$

The induction assumption says that there are no scalar term in $e'_{2r-3} e_l e_k e_k e_l$.

□

One can prove that $\mathrm{Cl}^{\mathbb{C}}(C)$ is a Hilbert space with the scalar product

$$\langle a|b\rangle = \mathrm{Tr}(a^* b). \tag{5.17}$$

Let $v \in V$ with $g(v,v) = 1$ (thus in $\mathrm{Cl}(V)$, we have $v^2 = 1$); since $v = \overline{v}$, we have

$$a^* v = vv^* = v^2 = 1.$$

Lemma 5.6.
The maps $a \mapsto ua$ and $a \mapsto au$ are unitary if and only if $uu^* = u^* u = 1$.

Proof. We pick $\lambda \in U(1)$ and $w = \lambda v \in V^{\mathbb{C}}$ which fulfils $w^* w = 1$. This is the most general element such that $ww^* = w^* w = 1$. Now for an arbitrary a, $b \in \mathrm{Cl}^{\mathbb{C}}(V)$, we compute the two followings:

$$\langle wa|wb\rangle = \mathrm{Tr}\left((wa)^* wb\right) = \mathrm{Tr}\left(a^* w^* wb\right) = \mathrm{Tr}(a^* b) = \langle a|b\rangle,$$

and

$$\langle aw|bw\rangle = \mathrm{Tr}\left(w^* a^* bw\right) = \mathrm{Tr}(ww^* a^* b) = \mathrm{Tr}(a^* b) = \langle a|b\rangle.$$

This proves that $a \mapsto wa$ and $a \mapsto aw$ are two unitary operators on the Hilbert space $\mathrm{Cl}^{\mathbb{C}}(V)$.

For the converse, we impose for all a, $b \in \mathrm{Cl}^{\mathbb{C}}(V)$:

$$\langle ua|ub\rangle = \mathrm{Tr}(ba^* u^* u) \overset{!}{=} \mathrm{Tr}(ba^*).$$

In particular with $a^* b = 1$, $\mathrm{Tr}(u^* u) = \mathrm{Tr}(1) = 1$, thus the scalar part of $u^* u$ is 1. So we write $u^* u = 1 + f$ where f is non scalar, and for any $x \in \mathrm{Cl}^{\mathbb{C}}(V)$, we have

$$\mathrm{Tr}(x) = \mathrm{Tr}(xu^* u) = \mathrm{Tr}(x) + \mathrm{Tr}(xf).$$

We conclude that $\mathrm{Tr}(xf) = 0$, and therefore that $f = 0$. □

5.3 Spinor representation

For the spinor representation, we restrict ourself to the even case $p + q = 2n$.

The aim of this subsection is to find some faithful representations of the complex Clifford algebra $\mathrm{Cl}^{\mathbb{C}}(p,q)$. In order to achieve this, we first consider $V^{\mathbb{C}}$, the complex vector space of V with an orthonormal basis $\{e_1, \cdots, e_{p-1}, e_p, \cdots, e_q\}$. The metric is $\eta(e_k, e_k) = 1$ and $\eta(e_{p+k}, e_{p+k}) = -1$ for $k = 0, \cdots, p-1$. We use the following basis:

$$f_k = \frac{1}{2}(e_k + e_{p+k}), \qquad g_k = \frac{1}{2}(e_k - e_{p+k}), \tag{5.18}$$

$$f_{p+s} = \frac{1}{2}(e_{2p+2s} + ie_{2p+2s+1}), \qquad g_{p+s} = \frac{1}{2}(e_{2p+2s} - ie_{2p+2s}) \tag{5.19}$$

for $k = 0, \cdots, p-1$. We note that $\{f_0, g_0\}$ spans a $\mathbb{C}^2$-space which is η-orthogonal to the one which is spanned by $\{f_1, g_1\}$. The following two spaces will prove to be useful:

$$W = \mathrm{Span}_{\mathbb{C}}\{f_0, f_1\} \simeq \mathbb{C}^2, \tag{5.20a}$$

$$\underline{W} = \mathrm{Span}_{\mathbb{C}}\{g_0, g_1\} \simeq \mathbb{C}^2. \tag{5.20b}$$

It is easy to compute the various products; among others we find

$$\eta(f_0, f_0) = 0, \quad \eta(f_1, f_0) = 0, \quad \eta(f_1, f_1) = 0; \tag{5.21}$$

so that for any $w \in W$, we have $\langle w, w\rangle = 0$; for this reason, we say that W is a **completely isotropic** subspace of $(V^{\mathbb{C}}, \eta^{\mathbb{C}})$. The space $\underline{W}$ has the same property.

Proposition 5.7.
We have

$$\underline{W} \simeq W^*, \tag{5.22}$$

where W^* is the dual space of W. By $\simeq$ we mean that there exists a linear bijective map $\psi\colon \underline{W} \to W^*$.

Proof. For each $\underline{w} \in \underline{W}$, we define $\psi(\underline{w}) \colon W \to \mathbb{C}$ by

$$\psi(\underline{w})(w) = \eta(w, \underline{w}).$$

We first show that the map ψ is injective. Let $\underline{w} \in \underline{W}$ be so that $\psi(\underline{w}) = 0$. Thus for all $v \in W$, we have

$$\psi(\underline{w})v = \eta(\underline{w}, v) = 0. \tag{5.23}$$

By decomposing $\underline{w} = ag_0 + bg_1$ and taking successively $v = f_0$ and $v = f_1$, we see that $a = b = 0$.

The next step is to see that the map ψ is surjective. We know that $dim_{\mathbb{C}}\underline{W} = dim_{\mathbb{C}}W^* = 2$ and that $\psi(g_0) \neq 0$. Let's prove that $\{\psi(g_0), \psi(g_1)\}$ is a basis of W^*. It is clear by linearity that $\{\psi(ag_0) : a \in \mathbb{C}\} = \text{Span}\{\psi(g_0)\}$. The fact that ψ is injective imposes that $\psi(g_1)$ doesn't belong to $\text{Span}\{\psi(g_0)\}$. So $\{\psi(g_0), \psi(g_1)\}$ is a two-dimensional free subset of W^*, and therefore is a basis of W^*. □

We turn our attention to the exterior algebra $\Lambda W = \mathbb{C} \oplus W \oplus (W \wedge W) \oplus \cdots \oplus \wedge^{p+q} W$ of W.

Definition 5.8.
We define the homomorphism $\tilde{\rho} \colon V^{\mathbb{C}} \to \text{End}(\Lambda W)$ by

$$\begin{aligned} \tilde{\rho}(f_i)\alpha &= f_i \wedge \alpha, \\ \tilde{\rho}(g_i)\alpha &= -\iota(g_i)\alpha \end{aligned} \tag{5.24}$$

($v \in V^{\mathbb{C}}$, $\alpha \in \Lambda W$) where ι denotes the interior product defined in page 19.

More explicitly, for all $z \in \mathbb{C}$ and for all $w, w' \in W$, we have

$$\tilde{\rho}(f_i)z = zf_i, \qquad \tilde{\rho}(g_i)z = 0, \tag{5.25a}$$
$$\tilde{\rho}(f_i)w = f_i \wedge w, \qquad \tilde{\rho}(g_i)w = -\eta(g_i, w)1, \tag{5.25b}$$
$$\tilde{\rho}(f_i)(w \wedge w') = 0, \qquad \tilde{\rho}(g_i)(w \wedge w') = -\eta(g_i, w)w' + \eta(g_i, w')w. \tag{5.25c}$$

We will see that, *via* some manipulations, $\tilde{\rho}$ provides a faithful representation of the Clifford algebra, the **spinor representation**.

Remark 5.9.
By "endomorphism of ΛW", we mean an endomorphism for the linear *structure of ΛW. We obviously not have $\tilde{\rho}(x)(\alpha \wedge \beta) = \tilde{\rho}(x)\alpha \wedge \tilde{\rho}(x)\beta$.*

Proposition 5.10.
The map $\tilde{\rho}$ is injective.

Proof. We have to show that $\tilde{\rho}(v) = 0$ (v in $V^{\mathbb{C}}$) implies $v = 0$. Any $v \in V^{\mathbb{C}}$ can be written as $v = a^i f_i + b^i g_i$ with a sum over i. We first have that

$$\tilde{\rho}(a^i f_i + b^i g_i)z = za^i f_i = 0,$$

but the f_i are independents and then $a^i = 0$. We can also write

$$\tilde{\rho}(b^0 g_0 + b^1 g_1)f_1 = -b^0\eta(g_0, f_1) - b^1\eta(g_1, f_1) = -\frac{b^1}{2} = 0,$$

then $b^1 = 0$. The same with f_0 proves that $b^0 = 0$. □

The homomorphism $\tilde{\rho}$ extends to the whole the tensor algebra of $V^{\mathbb{C}}$ by the following definitions:

$$\tilde{\rho}(1) = \text{id}_{\Lambda W}, \tag{5.26a}$$
$$\tilde{\rho}(e_k) = \tilde{\rho}(e_k), \tag{5.26b}$$
$$\tilde{\rho}(e_{k_1} \otimes \ldots \otimes e_{k_r}) = \tilde{\rho}(e_{k_1}) \circ \ldots \circ \tilde{\rho}(e_{k_r}). \tag{5.26c}$$

So we get $\tilde{\rho} \colon T(V^{\mathbb{C}}) \to \text{End}(\Lambda W)$. The following proposition will allow us to descent $\tilde{\rho}$ to a representation of the Clifford algebra.

Proposition 5.11.
The homomorphism $\tilde{\rho}$ maps $\mathcal{I}$ to 0: $\tilde{\rho}(\mathcal{I}) = 0$.

Problem and misunderstanding 21.
This proposition is wrong: there is a double covering.

Moreover, there is a sign problem in the proof: the sign in the first lines is not the one used in the definition of the Clifford algebra.

Proof. We have to check the following:

$$\tilde{\rho}(v \otimes w \oplus w \otimes v - 2\eta(v,w)1) = 0$$

for any choice of v, w in $\{e_0, e_1, e_2, e_3\}$. Here we will just check it explicitly for $v = e_0$ and $w = e_1$. The computation uses the definition (5.26c):

$$\begin{aligned}\tilde{\rho}(e_0 \otimes e_1 \oplus e_1 \otimes e_0 - 2\eta(e_0, e_1) &= \tilde{\rho}(e_0) \circ \tilde{\rho}(e_1) + \tilde{\rho}(e_1) \circ \tilde{\rho}(e_0) \\ &= 2\left[\tilde{\rho}(f_0)^2 - \tilde{\rho}(g_0)^2\right].\end{aligned} \tag{5.27}$$

It is easy to see that $\tilde{\rho}(f_0)^2 = 0$:

$$\tilde{\rho}(f_0)^2 \left[z \oplus w \oplus w_1 \wedge w_2\right] = \tilde{\rho}(f_0)[z f_0 \oplus f_0 \wedge w] = z f_0 \wedge f_0, = 0. \tag{5.28}$$

The proof that $\tilde{\rho}(g_0)^2 = 0$ is almost the same:

$$\tilde{\rho}(g_0)^2 \left[z \oplus w \oplus w_1 \wedge w_2\right] = \tilde{\rho}(g_0)[-\eta(g_0, w)1 \oplus -\eta(g_0, w_1) w_2 \oplus \eta(g_0, w_2) w_1].$$

□

We can now see $\tilde{\rho}$ as a map $\tilde{\rho}\colon \mathrm{Cl}^{\mathbb{C}}(p,q) \to \mathrm{End}(\Lambda W)$. By construction, it is a homomorphism and, thus, is a representation of $\mathrm{Cl}^{\mathbb{C}}(p,q)$ on ΛW. For compactness, we use the notation

$$\gamma_a := \sqrt{2}\tilde{\rho}(e_a). \tag{5.29}$$

Lemma 5.12.
The γ's operators satisfy the following relation:

$$\gamma_a \gamma_b + \gamma_b \gamma_a = -2\eta_{ab}\mathbb{1}. \tag{5.30}$$

Proof. We have to check this equality on any element of ΛW. If we choose $w_1 = a f_0 + b f_1$ and $w_2 = a' f_0 + b' f_1$, we find $w_1 \wedge w_2 = (ab' - ba') f_0 \wedge f_1$.

For example, we will explicitly check (5.30) with $a = b = 0$, i.e. $\tilde{\rho}(e_0) \circ \tilde{\rho}(e_0) = \frac{1}{2}\,\mathrm{id}$, which proves that $\gamma_0 \circ \gamma_0 = \mathrm{id}$.

$$\begin{aligned}\tilde{\rho}(e_0)^2[z \oplus w \oplus (ab' - ba') f_0 \wedge f_1] &= \tilde{\rho}(f_0 + g_0)^2[z \oplus w \oplus (ab' - ba') f_0 \wedge f_1] \\ &= \tilde{\rho}(f_0 + g_0)\Big[z f_0 \oplus f_0 \wedge w \oplus -\eta(g_0, w)1 \\ &\qquad - (ab' - ba')\eta(g_0, f_0) f_1 \\ &\qquad + (ab' - ba')\eta(g_0, f_1) f_0\Big] \\ &= \frac{1}{2}(z \oplus w \oplus (ab' - ba') f_0 \wedge f_1).\end{aligned} \tag{5.31}$$

□

Lemma 5.13.
For any sequence $i_0, \ldots i_3$ of 0 and 1 (with at least one of them equals to 1), we have

$$\mathrm{Tr}(\gamma_0^{i_0} \cdots \gamma_{2n-1}^{i_{2n-1}}) = 0. \tag{5.32}$$

We take the convention that $\gamma_a^0 = \mathbb{1}$.

Proof. If the number of nonzero i_k is even (say $2m$), we have:

$$\mathrm{Tr}(\gamma_{a_1} \ldots \gamma_{a_{2m}}) = \mathrm{Tr}(\gamma_{a_{2n}} \gamma_{a_1} \ldots \gamma_{a_{2m-1}})$$

because the trace is invariant under cyclic permutations. But we can also permute $\gamma_{a_{2m}}$ with the $2m-1$ other γ's. $\mathrm{Tr}(\gamma_{a_1} \ldots \gamma_{a_{2m}}) = (-1)^{2n-1}\,\mathrm{Tr}(\gamma_{a_{2m}} \gamma_{a_1} \ldots \gamma_{a_{2m-1}})$ because each permutation gives an extra minus sign (lemma 5.12). Then the trace is zero.

If the number of nonzero i_k is odd (say $2m-1$). Let $i_a = 0$ (we restrict ourself to the even dimensional case). We have $\mathrm{Tr}(A) = -\eta_{aa}\,\mathrm{Tr}(A\gamma_a\gamma_a)$. Using once again the cyclic invariance of the trace, $\mathrm{Tr}(\gamma_{a_1} \ldots \gamma_{a_{2m-1}} \gamma_a \gamma_a) = \mathrm{Tr}(\gamma_a \gamma_{a_1} \ldots \gamma_{a_{2m-1}} \gamma_a)$. But, if we permute the *first* γ_a with the $2m-1$ first γ's, we find $\mathrm{Tr}(\gamma_{a_1} \ldots \gamma_{a_{2m-1}} \gamma_a \gamma_a) = -\,\mathrm{Tr}(\gamma_a \gamma_{a_1} \ldots$ and the trace is zero again. □

Proposition 5.14.
The subset

$$\{\mathbb{1}, \gamma_a\gamma_b\,(a<b), \gamma_a\gamma_b\gamma_c\,(a<b<c), \cdots, \gamma_0\cdots\gamma_{2n}\}$$

is free in $\mathrm{End}(\Lambda W)$.

Proof. We consider a general linear combination of these operators:

$$E = \lambda\mathbb{1} + \sum_a \lambda_a\gamma_a + \sum_{a<b}\lambda_{ab}\gamma_a\gamma_b + \ldots + \sum_{a<b<c<d}\lambda_{abcd}\gamma_a\gamma_b\gamma_c\gamma_d.$$

The claim is that if $E=0$, then all the coefficients $\lambda_{(\ldots)}$ must be zero. First note that $Tr(E)=0=\lambda$ by lemma 5.13. It is also clear that $Tr(\gamma_i E)=0=\lambda_i$. In order to see that $\lambda_{ij}=0$, we compute $Tr(\gamma_j\gamma_i E)=0=\lambda_{ij}$. And so on. □

How many operators does we have in this free system ? Any operators in this system can be written as $\gamma_0^{i_0}, \cdots\gamma_{2n-1}^{i_{2n-1}}$ with i_k equal to zero or one. Thus we have 2^{2n} operators. On the other hand, we know that $dim_{\mathbb{C}}\Lambda W = 2p+2$, and then that $dim_{\mathbb{C}}\,\mathrm{End}(\Lambda W) = 4^2 = 16$. The result is that $\{\gamma_0^{i_0}, \cdots\gamma_{2n-1}^{i_{2n-1}}$ st $i_k = 0\,or\,1\}$ is a basis of $\mathrm{End}(\Lambda W)$. In other words (if we suppose a suitable ordering), the image by $\tilde\rho$ of

$$B = \{1, e_a, e_a\otimes e_b, e_a\otimes e_b\otimes e_c, e_a\otimes e_b\otimes e_c\otimes e_d\}$$

is a basis of $\mathrm{End}(\Lambda W)$.

If B is a basis of $C^{\mathbb{C}}_{(p,q)}$, then $\tilde\rho$ is bijective and thus isomorphic. Therefore, we expect $\tilde\rho\colon C^{\mathbb{C}}_{(p,q)} \to \mathrm{End}(\Lambda W)$ to be a faithful representation. It is not difficult to see that B is indeed a basis thanks to the equivalence relation.

5.3.1 Explicit representation

First, we choose a basis for ΛW:

$$1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad f_0 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad f_1 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad f_0\wedge f_1 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}. \tag{5.33}$$

Here is the explicit computation for the matrix γ_0 in this basis. First remark that $\tilde\rho(e_0)1 = f_0$, $\tilde\rho(e_0)f_0 = \frac{1}{2}$, $\tilde\rho(e_0)f_1 = f_0\wedge f_1$, $\tilde\rho(e_0)(f_0\wedge f_1) = \frac{1}{2}f_1$. Then

$$\begin{aligned}
\gamma_0\begin{pmatrix}1\\0\\0\\0\end{pmatrix} &= \sqrt{2}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, & \gamma_0\begin{pmatrix}0\\1\\0\\0\end{pmatrix} &= \sqrt{2}\begin{pmatrix}\frac{1}{2}\\0\\0\\0\end{pmatrix},\\
\gamma_0\begin{pmatrix}0\\0\\1\\0\end{pmatrix} &= \sqrt{2}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}, & \gamma_0\begin{pmatrix}1\\0\\0\\0\end{pmatrix} &= \sqrt{2}\begin{pmatrix}0\\0\\\frac{1}{2}\\0\end{pmatrix}.
\end{aligned} \tag{5.34}$$

This allows us to write down γ_0; the same computation gives the other matrices.

$$\begin{aligned}
\gamma_0 &= \sqrt{2}\begin{pmatrix}0&\frac{1}{2}&0&0\\1&0&0&0\\0&0&0&\frac{1}{2}\\0&0&1&0\end{pmatrix}, & \gamma_1 &= \sqrt{2}\begin{pmatrix}0&-\frac{1}{2}&0&0\\1&0&0&0\\0&0&0&-\frac{1}{2}\\0&0&1&0\end{pmatrix},\\
\gamma_2 &= \sqrt{2}\begin{pmatrix}0&0&-\frac{1}{2}&0\\0&0&0&\frac{1}{2}\\1&0&0&0\\0&-1&0&0\end{pmatrix}, & \gamma_3 &= \sqrt{2}\begin{pmatrix}0&0&-\frac{i}{2}&0\\0&0&0&\frac{i}{2}\\-i&0&0&0\\0&i&0&0\end{pmatrix}.
\end{aligned} \tag{5.35}$$

It is easy to check that these matrices satisfies (5.30).

Notice that, up to a suitable change of basis in ΛW, these are the usual Dirac matrices. Indeed we actually solved the physical problem to find a representation of the algebra (5.2). We understand by the way why do physicists work with 4-components spinors: the γ's are operators on the four-dimensional space ΛW; hence the Dirac operator will naturally acts on four-components objects.

The main result of this section is an explicit faithful representation of $\mathrm{Cl}^{\mathbb{C}}(p,q)$. This allows us to write a **Dirac operator** which solve (see the invitation 5.1 and [43]) the problem to find a "square root" of the d'Alembert operator: the differential operator $\mathcal{D} = \gamma^\mu\partial_\mu$ satisfies $\mathcal{D}^2 = \Box$.

5.3.2 A remark

Let us compare the two faithful representations

$$c\colon \operatorname{Cl}(V) \to \operatorname{End}_{\mathbb{R}}(\wedge V)$$
$$\tilde{\rho}\colon \operatorname{Cl}^{\mathbb{C}} \to \operatorname{End}_{\mathbb{R}}(\wedge W).$$

They obviously comes from the same ideas. One common point is that

$$c(e_1)(e_1 \wedge e_2) = 2\tilde{\rho}(e_1)(e_1 \wedge e_2) = e_2,$$

but they are different:

$$\tilde{\rho}(e_3)(e_0 \wedge e_2) = 0$$
$$c(e_3)(e_0 \wedge e_2) = e_3 \wedge e_0 \wedge e_1.$$

5.3.3 General two dimensional Clifford algebra

The Clifford algebra for the metric

$$g = \begin{pmatrix} \alpha & \delta \\ \delta & \beta \end{pmatrix}$$

is realised by matrices

$$\gamma_1 = \epsilon \begin{pmatrix} \sqrt{\alpha} & \\ & -\sqrt{\alpha} \end{pmatrix}, \quad \gamma_2 = \epsilon \begin{pmatrix} \delta/\sqrt{\alpha} & \beta - \delta^2/|\alpha| \\ 1 & -\delta/\sqrt{\alpha} \end{pmatrix}$$

where $\epsilon = \pm 1$ is chosen in such a way that $\epsilon|\alpha| = \alpha$.

5.4 Spin group

We will not immediately go on with Dirac operators on Riemannian manifolds because we still have to build some theory about the Clifford algebra itself. In particular, we have to define the spin group which will play a central role in the definition of the Dirac operator. Almost all –and (too ?) much more– the concepts we will introduce in this section can be found in [42]; a more physical oriented but useful approach can be found in [48].

Let define the map $\chi\colon \Gamma(p,q) \to GL(\mathbb{R}^{1,3})$ by

$$\chi(x)y = \alpha(x) \cdot y \cdot x^{-1}. \tag{5.36}$$

Let

$$\Gamma(p,q) = \{x \in \operatorname{Cl}(p,q) \text{ st } x \text{ is invertible and } \chi(x)y \in V \text{ for all } y \in V\}.$$

It should be remarked that this definition comes back to the real Clifford algebra. The Clifford algebra product gives this subset a group structure which is called the **Clifford group**. Any $x \in V$ is invertible since $x \cdot x = -\eta(x,x)1$, the inverse of x is given by $x^{-1} = x/\|x\|^2$.

The subset $\operatorname{Cl}(p,q)^+$ (resp. $\operatorname{Cl}(p,q)^-$) of $\operatorname{Cl}(p,q)$ is the image of even (resp. odd) tensors of $T(V)$ by the canonical projection $T(V) \to \operatorname{Cl}(p,q)$. With these definitions, we have a natural grading of Cl:

$$\operatorname{Cl}(p,q) = \operatorname{Cl}(p,q)^+ \oplus \operatorname{Cl}(p,q)^-, \tag{5.37}$$

and the subgroups

$$\Gamma(p,q)^+ = \Gamma(p,q) \cap \operatorname{Cl}(p,q)^+, \qquad \Gamma(p,q)^- = \Gamma(p,q) \cap \operatorname{Cl}(p,q)^-. \tag{5.38}$$

For $x_1, \ldots, x_n \in V$, we have $\tau(x_1 \cdots x_n) = x_n \cdots x_1$. The **spin group** is

$$\operatorname{Spin}(p,q) = \{x \in \Gamma(p,q)^+ | \tau(x) = x^{-1}\} \tag{5.39}$$

while the **spin norm** is the map $N\colon \Gamma(p,q) \to \Gamma(p,q)$ defined by

$$N(x) = x\tau(\alpha(x)).$$

We will see in proposition 5.25 that N actually takes its values in $\mathbb{R}$ and is therefore a homomorphism $N\colon \Gamma(p,q) \to \mathbb{R}$

Remark 5.15.
The elements of $\operatorname{Spin}(p,q)$ *are spin-normed at* 1*. Indeed, take a* s *in* $\operatorname{Spin}(p,q)$*. We have* $N(s) = s \cdot \tau(s) = 1$ *because* $\alpha(s) = s$ *and* $\tau(s) = s^{-1}$*. In particular* $\operatorname{Spin}(p,q) \cap \mathbb{R} = \mathbb{Z}_2$*.*

5.4.1 Studying the group structure

Proposition 5.16.
The set $\Gamma(p,q)$ admits a Lie group structure.

Proof. During this proof, μ denotes the Clifford multiplication: $\mu(x,y) = x \cdot y$. We know that $\mathrm{Cl}^{\mathbb{C}}(p,q)$ is isomorphic to $\mathrm{End}(\Lambda W)$ in which the multiplication is a continuous map. Thus μ is continuous on $\mathrm{Cl}^{\mathbb{C}}(p,q)$. But $\mathrm{Cl}(p,q)$ is a closed subset of $\mathrm{Cl}^{\mathbb{C}}(p,q)$, so μ is a continuous map in $\mathrm{Cl}(p,q)$. This proves that χ seen as a map from $\Gamma(p,q) \times V$ to V is a continuous map.

The space V is closed in $\mathrm{Cl}(p,q)$, thus $\sigma^{-1}(V)$ is also closed. But $\sigma^{-1}(V) = \Gamma(p,q) \times \mathrm{Cl}(p,q)$. So $\Gamma(p,q)$ is closed in $\mathrm{Cl}(p,q)$.

Now the result is just a consequence of theorems A.6 and **??**. Indeed, let us study the subset $\mathcal{I}$ which appears in the definitions of the Clifford algebra. It makes no difficult to convince ourself that it is a closed subgroup of $T(V)$. The theorem **??** thus makes $\mathrm{Cl}(p,q) = T(V)/\mathcal{I}$ a Lie group. But we just say that $\Gamma(p,q)$ is closed in $\mathrm{Cl}(p,q)$, and the fact that $\Gamma(p,q)$ is a subgroup of $\mathrm{Cl}(p,q)$ is clear. By theorem A.6 we conclude that there exists a Lie group structure on $\Gamma(p,q)$. □

Lemma 5.17.
The map χ is a homomorphism, in other words χ is a representation of $\Gamma(p,q)$.

Proof. The following computation uses the fact that α is a homomorphism:

$$\begin{aligned}\chi(a\cdot b)y &= \alpha(a\cdot b)\cdot y\cdot (a\cdot b)^{-1} = \alpha(a)\cdot\alpha(b)y\cdot b^{-1}\cdot a^{-1}\\ &= \alpha(a)\cdot\chi(b)y\cdot a^{-1} = \chi(a)\chi(b)y.\end{aligned}$$

□

Let $y \in \Gamma(p,q)^-$ and $v \in V$. Where is $y \cdot v$? First note that $(y\cdot v)^{-1} = v^{-1}\cdot y^{-1}$, so that

$$\begin{aligned}\alpha(y\cdot v)\cdot w\cdot (y\cdot v)^{-1} &= -\alpha(y)\cdot v\cdot w\cdot v^{-1}\cdot y^{-1}\\ &= -\alpha(y)\big(2\eta(v,w) - w\cdot v\big)\cdot v^{-1}\cdot y^{-1}\\ &= -2\eta(v,w)\alpha(y)\cdot v^{-1}\cdot y + \alpha(y)\cdot w\cdot y^{-1}\end{aligned} \tag{5.40}$$

which belongs to V because $y \in \Gamma(p,q)$. This reasoning shows that (apart for 0), $y\cdot v \in \Gamma(p,q)^+$ if and only if $y \in \Gamma(p,q)^-$.

Lemma 5.18.
If $x \in V$ is non-isotropic (i.e. $\eta(x,x) \neq 0$), the automorphism $\chi(x)$ is the orthogonal symmetry with respect to $x^\perp$.

We recall that

$$x^\perp = \{y \in V \text{ st } \eta(x,y) = 0\}.$$

We will denote by σ^x the orthogonal symmetry with respect to $x^\perp$.

Proof. When the operator σ^x acts on y, it just change the sign of the "x-part" of y. So we can write $\sigma^x y = y - 2\eta(x,y)1_x$, where $1_x := x/\|x\|$. It should be checked if $\chi(x)y = \alpha(x)\cdot y\cdot x^{-1}$ is equal to $y - 2\eta(x,y)1_x$ or not. We know that $x\cdot x = \eta(x,x)1 = -\|x\|$. It follows that

$$x\cdot y + y\cdot x = 2\eta(x,y)\frac{x\cdot x}{\|x\|}.$$

If we multiply this at right by x^{-1}, using the fact that $\alpha(x) = -x$, we find

$$-\alpha(x)\cdot y\cdot x^{-1} = -y + 2\eta(x,y)1_x,$$

which is precisely the identity we wanted to check. □

The following result will help us to identify subgroups of Clifford group as isometry groups.

Theorem 5.19 (Cartan-Dieudonné theorem)**.**
Each σ in $O(1,3)$ can be written as $\sigma = \tau_1 \circ \ldots \circ \tau_m$, where the τ's are orthogonal symmetries with respect to hyperplanes which are orthogonal to non-isotropic vectors.

Proposition 5.20.

$$\chi(\Gamma(p,q)) = O(p,q).$$

Proof. In order to show that $\chi(\Gamma(p,q)) \subset O(p,q)$ take $z \in V$ and $x \in \Gamma(p,q)$. Since $\alpha(x)\cdot z\cdot x^{-1}$ lies in V, we can write:

$$\alpha(x)\cdot z\cdot x^{-1} = -\alpha\left(\alpha(x)\cdot z\cdot x^{-1}\right) = -x\cdot\alpha(z)\cdot\alpha(x^{-1}) = x\cdot z\cdot\alpha(x^{-1}).$$

In order to see that $\chi(x) \in O(p,q)$, we have to prove that $\|\chi(x)y\|^2_{(p,q)} = \|y\|^2_{(p,q)}$. This is achieved by the following computation:

$$\begin{aligned}\|\chi(x)y\|^2_{(p,q)} &= -\left(\alpha(x)\cdot y\cdot x^{-1}\right)^2 = \left(\alpha(x)\cdot y\cdot x^{-1}\right)\left(x\cdot y\cdot\alpha(x^{-1})\right)\\ &= -\alpha(x)\cdot y^2\cdot\alpha(x^{-1}) = \|y\|^2_{(p,q)}.\end{aligned} \tag{5.41}$$

The last step is simply the fact that $y^2 \in \mathbb{R}$ and therefore commutes with anything. We now know that $\chi(x) \in O(p,q)$ for all $x \in \Gamma(p,q)$. Thus $\chi(\Gamma(p,q)) \subset O(p,q)$.

For the second part, let σ be in $O(p,q)$. The Cartan-Dieudonné theorem(theorem 5.19) says that $\sigma = \sigma^{x_1}\circ\ldots\circ\sigma^{x_r}$ for some $x_1,\ldots,x_r$ in V. Thus $\sigma = \chi(x_1\cdots x_r)$, and $O(p,q) \subset \chi(\Gamma(p,q))$. □

Proposition 5.21.

$$\ker\chi = \mathbb{R}^\times \tag{5.42}$$

where the right hand side is the set of invertible elements of $\mathbb{R}$.

Proof. Before beginning the proof, we want to insist on the fact that $x \in \ker\chi$ does not mean that $\chi(x)y = 0$ for all y in V. The "zero" of an algebra is the element e which satisfies $e\cdot y = y\cdot e = y$ for all y in the algebra. In other words, x is in the kernel of χ if and only if $\chi(x) = \mathrm{id}$.

First we show that $\mathbb{R}_0 \subset \ker\chi$. If $x \in \mathbb{R}$, then $\alpha(x) = x$. Therefore, when $x \neq 0$,

$$\chi(x)y = \alpha(x)\cdot y\cdot x^{-1} = y,$$

because the algebra product $\cdot$ between an element of $\mathrm{Cl}(p,q)$ and a real is commutative. Note that this does not work with $x = 0$.

We are now going to show that $\ker\chi \subset \mathbb{R}$. Let $z \in \ker\chi$. We decompose (definitions (5.38)) it into his odd and even part: $z = z^+ + z^-$, with $z^\pm \in \Gamma(p,q)^\pm$. These two can be written as $z^+ = e_{j_1}\cdots e_{j_{2r}}$ and $z^- = e_{i_1}\cdots e_{i_{2r-1}}$ with no two i_k or j_k equals. This is almost the general form of elements in even and odd part of $\Gamma(p,q)$: the only other possibility is z in $\mathbb{R}$. Obviously $\alpha(z^\pm) = \pm z^\pm$. Multiplying the condition $\chi(z)y = y$ at right by $(z^+ + z^-)$, we find

$$(z^+ - z^-)y = y(z^+ + z^-).$$

Thanks to equation (5.37), we can split this condition into even and odd parts:

$$z^+y = yz^+, \qquad\qquad z^-y = -yz^-. \tag{5.43}$$

The first equation with $y = e_{j_1}$ gives $e_{j_1}\cdots e_{j_{2r}}\cdot e_{j_1} = e_{j_1}e_{j_1}\cdots e_{j_{2r}}$. In the left hand side, permute the last e_{j_1} from last to second position. So we find the right hand side, with an extra minus sign. This means that $z^+ = 0$. In the same way, the second equation gives $z^- = 0$. We are left with the last possibility: $z \in \mathbb{R}$. □

Corollary 5.22.
For any $s \in \Gamma(p,q)$, there exists some non-isotropic vectors $x_1,\ldots,x_r$, and $c \in \mathbb{R}$ such that $s = cx_1\cdots x_r$.

Proof. Let us take a $s \in \Gamma(p,q)$; we just saw (theorem 5.20) that $\chi(s)$ is an element of $O(p,q)$. It can be written $\chi(s) = \sigma_1\circ\ldots\circ\sigma_m$. But we had shown that $\sigma_i = \chi(x_i)$ for any x_i normal to the hyperplane defining σ_i. We thus have

$$\chi(s) = \chi(x_1\cdots x_m),$$

where s belongs to $\Gamma(p,q)$ and is therefore invertible. This leads us to write $\mathrm{id} = \chi(s^{-1}\cdot x_1\cdots x_m)$. But the kernel of χ is $\mathbb{R}$ (proposition 5.21); so one can find a $r \in \mathbb{R}$ such that $s^{-1}\cdot x_1\cdots x_m = r$. The claim follows. □

Lemma 5.23.
If $v \in V$,

$$\det\chi(v) = -1. \tag{5.44}$$

Proof. We already know that $det\chi(v) = \pm 1$. To check that the right sign is plus, take the following basis of V: $\{v, v_i^\perp\}$ where $\{v_i^\perp\}$ is a basis of $v^\perp$. Calculating the action of $\chi(v)$ on this basis, we find:

$$\begin{aligned} \chi(v)v &= -v \cdot v \cdot v^{-1} = -v, \\ \chi(v)v_i^\perp &= -v \cdot v_i^\perp \cdot v^{-1} = v_i^\perp \cdot v \cdot v^{-1} = v_i^\perp. \end{aligned} \tag{5.45}$$

In this computation, we used the relation $v \cdot w = -w \cdot v - 2\langle v, w\rangle$ which is true for all v, w in V. The action of $\chi(v)$ on this basis is thus to let unchanged three vectors and to change the sign of the fourth. This proves the claim. □

Theorem 5.24.

$$\chi(\Gamma(p,q)^+) = \mathrm{SO}(p,q). \tag{5.46}$$

Proof. From corollary 5.22, and definition 5.38, an element $s \in \Gamma(p,q)^+$ reads $s = cv_1 \cdots v_{2r}$. Thus

$$\det\chi(s) = \det\chi(v_1 \cdots v_{2r}) = \det\left[\chi(v_1)\ldots\chi(v_{2r})\right]. \tag{5.47}$$

But we know that, for all v_i in V, $det\chi(v_i) = -1$. So $\det\chi(s) = 1$ and $\chi(\Gamma(p,q)^+) \subseteq \mathrm{SO}(p,q)$. As set,

$$\Gamma(p,q) = \Gamma(p,q)^+ \cup \Gamma(p,q)^-,$$

but the lemma shows that $\det\chi(\Gamma(p,q)^-) = -1$ so, from theorem 5.20, $\chi(\Gamma(p,q)^+)$ must be the whole $\mathrm{SO}(p,q)$. □

Proposition 5.25.
The map N takes values in $\mathbb{R}$ and the formula

$$N(x \cdot y) = N(x)N(y), \tag{5.48}$$

holds for all x, $y \in \Gamma(p,q)$.

Proof. We write as usual $x \in \Gamma(p,q)$ as $x = cv_1 \cdots v_r$. So,

$$N(x) = cv_1 \cdots v_r\tau(\alpha(cv_1 \cdots v_r)) = (-1)^r c^2 v_1 \cdots v_r \cdot v_r \cdots v_1. \tag{5.49}$$

The first equality comes from the fact that $\alpha(cv_1 \cdots v_r) = (-1)^r cv_1 \cdots v_r$. Now $N(x) \in \mathbb{R}$ because $v_i \cdot v_i = -\langle v_i, v_i\rangle \in \mathbb{R}$ for all i. Hence the following hold:

$$\begin{aligned} N(x \cdot y) &= v \cdot y \cdot \tau(\alpha(v \cdot y)) \\ &= v \cdot y \cdot \tau(\alpha(y)) \cdot \tau(\alpha(v)) \\ &= v \cdot N(y)\tau(\alpha(v)) \\ &= N(y)N(x). \end{aligned} \tag{5.50}$$

This is the claim. □

Theorem 5.26.
We have the following isomorphism of groups

$$\mathrm{Spin}(p,q) = \mathrm{SO}_0(p,q).$$

provided by the map χ.

$\boxed{\textbf{Problem and misunderstanding}}$ **22.**
This result is wrong because of a double covering issue. The real proposition is the next one. I should try to merge the proofs.

Proof. Let $\{e_1, \cdots, e_p, f_1, \cdots, f_p\}$ be a basis of $\mathbb{R}^{p+q}$ where the e_i's are time-like and the f_j's are space-like. Following the discussion at page **??**, we have

$$\mathrm{SO}(p,q) = \mathrm{SO}_0(p,q) \cup \xi\,\mathrm{SO}_0(p,q)$$

where ξ is defined as follows: $\xi e_1 = -e_1$, $\xi f_1 = -f_1$ and $\xi e_k = e_k$, $\xi f_k = f_k$ for $k \neq 1$. This element can be implemented as $\xi = \chi(g)$ for $g = e_1 f_1$. It is easy to see that $g^{-1} = -f_1 e_1$ and that $\tau(g) = f_1 e_1$, so that $g \notin \mathrm{Spin}(p,q)$.

Is it possible to find another $h \in \Gamma(p,q)$ such that $\chi(h) = \xi$? If $\chi(a) = \chi(b)$ for $a, b \in \Gamma(p,q)$, then $a = rb$ for a certain $r \in \mathbb{R}$. So we find that $h = g^{-1}/r$ is the general form of an element in $\Gamma(p,q)$ such that $\chi(h) = \xi$. This is an element of $\mathrm{Spin}(p,q)$ if and only if $\tau(h) = h^{-1}$, or $-e_1 f_1/r = re_1 f_1$ which has no solutions. We conclude that no element of $\mathrm{Spin}(p,q)$ is send on ξ by χ. So

$$\chi\big(\mathrm{Spin}(p,q)\big) \subset SO_0(p,q).$$

$\boxed{\textbf{Problem and misunderstanding}}$ **23.**
I still have to prove the surjectivity of χ from $\mathrm{Spin}(p,q)$ to $\mathrm{SO}(p,q)$.

$\square$

Theorem 5.27.

$$\chi(\mathrm{Spin}(p,q)) = \mathrm{SO}_0(p,q) \tag{5.51}$$

where the index 0 means the identity component.

Proof. Proposition 5.21, theorem 5.24 and remark 5.15 show that the map $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}(p,q)$ is a homomorphism with $\mathbb{Z}_2$ as kernel. We begin to prove that $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}_0(p,q)$ is surjective. On the one hand, elements of $\mathrm{Spin}(p,q)$ satisfy one more condition than the ones of $\Gamma(p,q)^+$. Thus the algebra $\mathrm{Spin}(p,q)$ has codimension one in $\Gamma(p,q)^+$.

On the other hand, we know that $\mathrm{SO}(p,q) = \mathrm{SO}_0(p,q) \cup h\,\mathrm{SO}_0(p,q)$ where h is the matrix such that $he_i = -e_i$ for $i = 0,\ldots,3$. Since $\mathrm{Spin}(p,q)$ has codimension one in $\Gamma(p,q)^+$, there is at most one more generator in $\chi(\Gamma(p,q)^+)$ than in $\chi(\mathrm{Spin}(p,q))$ (because χ is a homomorphism). In order to prove this theorem, we just need to show that elements of $\chi(\Gamma(p,q)^+)$ which do not belong to $\chi(\mathrm{Spin}(p,q))$ is h.

Is is no difficult to see that $\chi(e_0\cdot e_1\cdot e_2\cdot e_3)e_i = -e_i$ for $i = 0\ldots 3$: just write $\chi(e_0\cdot e_1\cdot e_2\cdot e_3)e_i = e_0\cdot e_1\cdot e_2\cdot e_3\cdot e_i\cdot e_3^{-1}\cdot e_2^{-1}\cdot e_1^{-1}\cdot e_0^{-1}$ and use the commutation relations. An easy computation gives $N(e_0\cdot e_1\cdot e_2\cdot e_3) = -1$; then this is not in $\mathrm{Spin}(p,q)$ by remark 5.15. $\square$

We write it by the exact sequence

$$\mathbb{Z}_2 \hookrightarrow \mathrm{Spin}(p,q) \xrightarrow{\chi} \mathrm{SO}_0(p,q) \tag{5.52}$$

we say that the group $\mathrm{Spin}(p,q)$ is a **double covering** of $\mathrm{SO}_0(p,q)$.

Lemma 5.28.
If $\pi\colon \tilde{X} \to X$ is a covering which satisfies

(i) X is path connected,

(ii) $\forall x \in X$, $\tilde{X}_x := \pi^{-1}(x)$ is path connected in $\tilde{X}$ i.e. *for all $a, b \in \tilde{X}$, there exist a path in $\tilde{X}$ which joins a and b,*

then $\tilde{X}$ is path connected.

Proof. If $\tilde{x}$ and $\tilde{y}$ are in $\tilde{X}$, we can suppose that $\pi(\tilde{x}) \neq \pi(\tilde{y})$ (because if $\pi(\tilde{x}) = \pi(\tilde{y})$, the second assumption gives the thesis). We define x and y as their projections: $x = \pi(\tilde{x})$ and $y = \pi(\tilde{y})$. Let γ be a path such that $\gamma(0) = x$ and $\gamma(1) = y$, and $\tilde{\gamma}$ be the lift of γ in $\tilde{X}$ which contains $\tilde{x}$: $\tilde{\gamma}(0) = \tilde{x}$ and $\pi(\tilde{\gamma}(1)) = \gamma(1) = y$. Then $\tilde{\gamma}(1)$ lies in $\tilde{X}_y$. Therefore, we can consider γ' which joins $\tilde{\gamma}(1)$ and $\tilde{y}$.

So, $\gamma' \circ \tilde{\gamma}$ is a path which contains $\tilde{x}$ and $\tilde{y}$. $\square$

Proposition 5.29.
The group $\mathrm{Spin}(p,q)$ is connected.

Proof. We will prove that the covering $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}_0(p,q)$ fulfils lemma 5.28. We just have to show that $\mathrm{Spin}(p,q)$ fulfills the second assumption of the lemma. First note that $\chi(\tilde{x}) = \chi(\tilde{y})$ implies $\chi(\tilde{x}\tilde{y}^{-1}) = e$, and then $\tilde{x} = \pm\tilde{y}$ because of proposition 5.21. Since the other case is trivial, we can suppose $\tilde{x} = -\tilde{y}$.

It remains to prove that for every $g \in \mathrm{Spin}(p,q)$, there is a path in $\mathrm{Spin}(p,q)$ which joins g and $-g$. The answer is given by the path $t \mapsto \gamma(t)g$ where

$$\gamma(t) = \exp(te_1\cdot e_2) = \cos(t)(-1) + \sin(t)e_1\cdot e_2$$

which satisfies $\gamma(0) = 1$ and $\gamma(\pi) = -1$. $\square$

Proposition 5.30.
The homomorphism $\tilde{\rho}$ restricts to a homomorphism $\tilde{\rho}\colon \mathrm{Spin}(p,q) \to \mathrm{GL}(\Lambda^+ W)$.

Proof. An element in $\mathrm{Spin}(p,q)$ reads $s = cv_1 \cdots v_{2r}$ and its image by $\tilde{\rho}$ is

$$\tilde{\rho}(s) = c\tilde{\rho}(v_1) \circ \cdots \circ \tilde{\rho}(v_{2r}).$$

When one applies $\tilde{\rho}(v_1)$ to an element $\alpha \in \Lambda^k W$, one obtains a linear combination of an element of $\Lambda^{k-1}W$ and one of $\Lambda^{k+1}W$. The element $\tilde{\rho}(s)$ being an even composition of such maps, its transforms an element of $\Lambda^+ W$ into an element of $\Lambda^+ W$. □

Notice that an element of V —no $V^{\mathbb{C}}$— is represented on $\Lambda^+ W$ by complex matrices. This is not a problem. In the case of $\mathbb{R}^{1,3}$, we have $\dim \Lambda^+ W = 2$ and thus

$$\tilde{\rho}\big(\mathrm{Spin}(1,3)\big) \subset \mathrm{GL}(2,\mathbb{C}).$$

The following is the lemma 8.5 (page 57) of [37].

Lemma 5.31.
Let $\rho\colon \mathrm{Cl}(p,q) \to \mathrm{Hom}_{\mathbb{C}}(E,E)$ be a representation of the Clifford algebra on a vector space E. If $p+q \geqslant 2$, then for all $s \in \mathrm{Spin}(p-1,q) \subset \mathrm{Cl}(p,q)$,

$$\det{}_{\mathbb{C}}\big(\rho(s)\big) = \pm 1.$$

Proof. No proof. □

Theorem 5.32.
The representation $\tilde{\rho}$ provides a group isomorphism

$$\mathrm{Spin}(1,3) \simeq \mathrm{SL}(2,\mathbb{C})$$

Proof. In the case $p = 2$, $q = 3$, the lemma assures us that for each s in the spin group, $\det \tilde{\rho}(s) = 1$. Since $\mathrm{Spin}(1,3)$ is connected and the determinant function is continuous, we deduce that $\det \tilde{\rho}(s) \equiv 1$. This proves that $\tilde{\rho}\big(\mathrm{Spin}(1,3)\big) \subset \mathrm{SL}(2,\mathbb{C})$. The proposition 2.1 thus implies that

$$\tilde{\rho}\big(\mathrm{Spin}(1,3)\big) = \mathrm{SL}(2,\mathbb{C}),$$

but from $\mathrm{Cl}(1,3)$, the representation $\tilde{\rho}$ is yet injective. *A forciori*, the representation $\tilde{\rho}$ is injective from $\mathrm{Spin}(1,3)$. This finishes the proof. □

5.4.2 Redefinition of $\mathrm{Spin}(V)$

As it, this new definition only holds when g is positive defined.

Problem and misunderstanding 24.
When we work with a signature (p,q), maybe we only get the connected part. To be checked.

Let us take v, $x \in V$ with $g(v,v) = 1$. We have

$$-vxv^{-1} = -vxv = -2g(x,v)v + xv^2 = x - 2g(x,v)v \in V.$$

The effect was to reverse the v component of x; the map $x \mapsto -vxv^{-1}$ is σ^v. Now, when $\lambda \in U(1)$ and $w = \lambda v$, we also have that $x \mapsto -wxw^{-1}$ is σ^v. Now we look at $\chi(a)\colon x \mapsto \alpha(a)xa^{-1}$ with $a = w_1 \ldots w_r$, a product of unitary vectors in $V^{\mathbb{C}}$. Explicitly,

$$\chi(a)x = (-1)^r w_1 \ldots w_r x w_r^{-1} \ldots w_1^{-1},$$

a composition of reflexions in V. When r is even, it is a rotation. We conclude that when a is an even product of unitary vectors in $V^{\mathbb{C}}$, then $\chi(a) \in \mathrm{SO}(V)$. Theorem 5.19 states that any rotation of V is a composition of reflexions. So we define

$$\mathrm{Spin}^c(V) = \{w_1 \ldots w_{2k} \text{ st } w_j \in V^{\mathbb{C}},\ w_j^* w_j = 1\} \subset {\mathrm{Cl}^{\mathbb{C}}}^0(V), \tag{5.53}$$

and $\chi\colon \mathrm{Spin}^c(V) \to \mathrm{SO}(V)$ is a surjective group homomorphism. The inverse in $\mathrm{Spin}^c(V)$ is given by

$$(w_1 \ldots w_{2k})^{-1} = w_{2k}^* \ldots w_1^* = \overline{w_{2k}} \ldots \overline{w_1}.$$

In the real case, proposition 5.21 says that $\ker\chi = \mathbb{R}^\times$. In the complex case we get $\ker\chi = \mathbb{C}^\times$ and, when we look at $\ker\chi|_{\mathrm{Spin}^c(V)}$, we find

$$\ker\chi = U(1). \tag{5.54}$$

Then we find the short exact sequence

$$1 \xrightarrow{\text{id}} U(1) \xrightarrow{\text{id}} \mathrm{Spin}^c(V) \xrightarrow{\chi} \mathrm{SO}(V) \xrightarrow{\text{id}} 1. \tag{5.55}$$

Let $u = w_1 \ldots w_{2k} \in \mathrm{Spin}^c(V)$ with $w_j = \lambda_j v_j$ and $\lambda_j \in V$, so $\tau(u) = w_{2k} \ldots w_1$ and

$$\tau(u)u = w_{2k} \ldots w_1 w_1 \ldots w_{2k} = \lambda_1^2 \ldots \lambda_{2k}^2 \in U(1).$$

This proves that $\tau(u)u$ is central in $\mathrm{Spin}^c(V)$. We define the homomorphism

$$\begin{aligned} \nu\colon \mathrm{Spin}^c(V) &\to U(1) \\ u &\mapsto \tau(u)u. \end{aligned} \tag{5.56}$$

This is a homomorphism because

$$\begin{aligned} \nu(u_1u_2) &= \tau(u_1u_2)u_1u_2 = \tau(u_2)\underbrace{\tau(u_1)u_1}_{\text{central}}u_2 = \tau(u_2)u_2\tau(u_1)u_1 \\ &= \nu(u_2)\nu(u_1) = \nu(u_1)\nu(u_2). \end{aligned}$$

The map ν naturally restricts to $U(1)$ as

$$\nu(\lambda) = \lambda^2.$$

The combined map $(\chi,\nu)\colon \mathrm{Spin}^c(V) \to \mathrm{SO}(V) \times U(1)$ has kernel $\{\pm 1\}$. We define

$$\mathrm{Spin}(V) = \ker\nu|_{\mathrm{Spin}^c(V)}. \tag{5.57}$$

Lemma 5.33.
This group is the same as the one defined in equation (5.39).

Proof. Let $u \in \mathrm{Spin}(V)$ (in the sense of equation (5.57)). The fact for u to belongs to $\mathrm{Spin}(V)$ implies the two following:

(i) $u \in \mathrm{Spin}^c(V) \Rightarrow u^*u = 1$,

(ii) $u \in \ker\nu \Rightarrow \tau(u)u = 1$.

The second point says that $u^{-1} = \tau(u)$, which is a first good point to fit the first definition of $\mathrm{Spin}(V)$. Now we have to prove that $u \in \Gamma^+(V)$: u must be invertible and $\chi(u)x$ must belongs to V for all $x \in V$. These two points are contained in the definition of $\mathrm{Spin}^c(V)$. □

Let us see in the new definition how is $\chi\colon \mathrm{Spin}(V) \to \mathrm{SO}(V)$. On $\mathrm{Spin}^c(V)$, we have $\ker\chi = U(1)$, but on $\mathrm{Spin}(V)$ we require moreover $\tau(u)u = 1$, thus an element of $\ker\chi$ in $\mathrm{Spin}(V)$ fulfils $\tau(\lambda)\lambda = 1$, so that $\lambda = \{\pm 1\}$. We conclude that $\ker\chi|_{\mathrm{Spin}(V)} = \{\pm 1\}$, and then that $\mathrm{Spin}(V)$ is a double covering of $\mathrm{SO}(V)$.

5.4.3 A few about Lie algebra

Proposition 5.34.
We have an isomorphism

$$\mathfrak{spin}(p,q) \simeq \mathfrak{so}(p,q)$$

between the Lie algebras of $\mathrm{Spin}(p,q)$ *and* $\mathrm{SO}(p,q)$.

Proof. Using the second part of lemma A.7, with the map $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}(p,q)$, we find that $d\chi_e(\mathfrak{spin}(p,q)) = \mathfrak{so}(p,q)$. Then we know (lemma A.8) that

$$\mathfrak{so}(p,q) = \mathfrak{spin}(p,q)/\ker d\chi_e.$$

On the other hand, the first part of the same lemma gives us that $\chi^{-1}(e)$ is a Lie subgroup of $\mathrm{Spin}(p,q)$ whose Lie algebra is $\ker d\chi_e$. But $\chi^{-1}(e) = \mathbb{Z}_2$, so $\ker d\chi_e = \{0\}$. □

Let us now shortly speak about the Lie algebra of $\Gamma(p,q)^+$. A basis of $\mathrm{Cl}(p,q)^+$ is

$$\{1, \gamma_0\cdot\gamma_1, \gamma_0\cdot\gamma_1, \gamma_0\cdot\gamma_3, \gamma_0\cdot\gamma_1\cdot\gamma_2\cdot\gamma_3\}.$$

Thanks to the anticommutation relations, we don't need $\gamma_1\cdot\gamma_2$ in the basis.

Remember that $\Gamma(p,q)^+$ is the set of the $x\in\mathrm{Cl}^+(p,q)$ such that $x\cdot v\cdot\alpha(x^{-1})$ lies in V for all $v\in V$. Let $x(t)$ be a path in $\Gamma(p,q)^+$ such that $x(0)=e$ and $\dot x(0)=X$. Differentiating the definition relation, we find

$$\dot x\cdot v\cdot\alpha(x^{-1})|_0 + x\cdot v\cdot(-)\alpha(\dot x)|_0 = X\cdot v - v\cdot X,$$

therefore

$$\mathfrak{Lie}(\Gamma(p,q)^+) = \left\{X\in\mathrm{Cl}^+(p,q)\text{ such that }X\cdot v - v\cdot X\in V,\ \forall v\in V\right\}.$$

It is clear that $\mathbb{C}$ is a subset of $\mathfrak{Lie}(\Gamma(p,q)^+)$, and that V is not. The following computation shows that $V\cdot V$ is a subset $\mathfrak{Lie}(\Gamma(p,q)^+)$:

$$a\cdot b\cdot v - v\cdot a\cdot b = 2\eta(v,a)b - 2\eta(v,b)a.$$

We can also check that $V\cdot V\cdot V\cdot V\cap\mathfrak{Lie}(\Gamma(p,q)^+)=\varnothing$. A basis of $\mathfrak{Lie}(\Gamma(p,q)^+)$ is

$$\{1, e_\alpha\cdot e_\beta\text{ st }\alpha<\beta\}$$

We know that $\ker[\chi\colon\Gamma(p,q)^+\to\mathrm{SO}(p,q)]=\mathbb{R}_0$. So the kernel of the restriction of $d\chi_e$ to $\mathfrak{Lie}(\Gamma(p,q)^+)$ is the Lie algebra of $\mathbb{R}_0$ (see lemma A.7), which is $\mathbb{R}$. Therefore, a basis of $\mathfrak{spin}(p,q)$ is

$$\{e_\alpha\cdot e_\beta\text{ st }\alpha<\beta\}.$$

5.4.4 Grading ΛW

We already know that $\Lambda W=\mathbb{C}\oplus W\oplus\Lambda^2W$. This space can be written as

$$\Lambda W = \Lambda W^+\oplus\Lambda W^-,$$

with $\Lambda W^+=W$ and $\Lambda W^-=\mathbb{C}\oplus\Lambda^2W$. The interest of such a decomposition lies in the definition of an action of $\mathrm{Cl}^+(p,q)$ on ΛW. This action will be defined by $\bullet\colon\mathrm{Cl}^+(p,q)\times\Lambda W\to\Lambda W$,

$$x\bullet\alpha = \tilde\rho(x)\alpha$$

for any x in $\mathrm{Cl}^+(p,q)$ and any α in ΛW (see definition 5.8).

Proposition 5.35.
This action preserves the grading of ΛW:

$$\begin{aligned}\mathrm{Cl}^+(p,q)\bullet\Lambda W^+ &= \Lambda W^+\\ \mathrm{Cl}^+(p,q)\bullet\Lambda W^- &= \Lambda W^-.\end{aligned}\tag{5.58}$$

Proof. For $x\in\mathbb{C}$, theses equalities are obvious. We have to check it for $x=e_i\cdot e_j$. Here, we will just check that $(e_1\cdot e_0)\bullet(v\wedge w)\in\Lambda W^+$. This follows from a simple computation:

$$\begin{aligned}\tilde\rho(e_1)\tilde\rho(f_0+g_0)(v\wedge w) &= \tilde\rho(f_1+g_1)\left[-\eta(g_0,v)w+\eta(g_0,w)v\right]\\ &= -\eta(g_0,v)f_1\wedge w+\eta(g_0,w)f_1\wedge v\\ &\quad+\eta(g_0,v)\eta(g_1,w)-\eta(g_0,w)\eta(g_1,v).\end{aligned}\tag{5.59}$$

□

Since $\mathrm{Spin}(p,q)$ is a subgroup of $\mathrm{Cl}^+(p,q)$, we can construct two new representation of $\mathrm{Spin}(p,q)$. These are $\rho^\pm\colon\mathrm{Spin}(p,q)\times\Lambda W^\pm\to\Lambda W^\pm$,

$$\begin{aligned}\rho^-(s)w^- &= \tilde\rho(s)w^-,\\ \rho^+(s)w^+ &= \tilde\rho(s)w^+,\end{aligned}\tag{5.60}$$

for $w^\pm$ in $\Lambda W^\pm$. This is no more than the fact that $\tilde\rho$ is reducible and that two invariant subspaces are ΛW^+ and ΛW^-.

5.4.5 Clifford algebra for $V = \mathbb{R}^2$

5.4.5.1 General definitions

The whole construction can also be applied to $V = \mathbb{R}^2$ with the Euclidean metric. This is our business now. We take the complex vector space $V^{\mathbb{C}}$ and an orthonormal basis $\{e_1, e_2\}$. As before, we define

$$f_1 = \frac{1}{2}(e_1 + ie_2), \qquad g_1 = \frac{1}{2}(e_1 - ie_2).$$

There are no difficulties to see that $Span(f_1)$ is a completely isotropic subspace of $V^{\mathbb{C}}$. Thus we define $W = \mathbb{C}f_1$, $\Lambda W = \mathbb{C} \oplus W$, $\Lambda W^+ = \mathbb{C}$, and $\Lambda W^- = W$. The homomorphism $\tilde{\rho}\colon V^{\mathbb{C}} \to \mathrm{End}(\Lambda W)$ in ΛW is defined by

$$\begin{aligned} \tilde{\rho}(f_1)\alpha &= f_1 \wedge \alpha, \\ \tilde{\rho}(g_1)\alpha &= -i(g_1)\alpha, \end{aligned} \tag{5.61}$$

where α is any element of ΛW. In the basis $1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $f_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, we easily find that

$$\tilde{\rho}(e_1) = \begin{pmatrix} 0 & -\frac{1}{2} \\ 1 & 0 \end{pmatrix}, \quad \tilde{\rho}(e_2) = \begin{pmatrix} 0 & -\frac{i}{2} \\ -i & 0 \end{pmatrix}.$$

For $c \in \mathbb{R}$ we also have $\tilde{\rho}(c)f_1 = cf_1$ and $\tilde{\rho}(c)1 = c$, thus we assign the matrix $\begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}$ to the number c.

As before, we define $\gamma_i = \sqrt{2}\tilde{\rho}(e_i)$. We immediately have $\gamma_1\gamma_2 + \gamma_2\gamma_1 = 0$ and $\gamma_i\gamma_i = -2\mathbb{1}$, so that the γ's satisfy the Clifford algebra for the euclidian metric.

For notational conveniences, it proves useful to make a change of basis so that we get

$$\gamma_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_2 = -\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{5.62}$$

The algebra Cl(2) is isomorphic to the algebra which is generated by direct sum $\mathrm{Cl}(2) \simeq \mathbb{R} \oplus \gamma_1 \oplus \gamma_2 \oplus \mathbb{R}\gamma_1\gamma_2$. A general element of Cl(2) can be written as $x\gamma_1 + y\gamma_2 + x'\mathbb{R} + y'\gamma_1\gamma_2$. In the representation of $\tilde{\rho}$, a general element of Cl(2) is therefore

$$\begin{pmatrix} x' + iy' & x + iy \\ -x + iy & x' - iy' \end{pmatrix},$$

so that we can write the Clifford algebra of $\mathbb{R}^2$ as

$$\mathrm{Cl}(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \,:\, \alpha, \beta \in \mathbb{C} \right\}.$$

The following four matrices provide a basis:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad i = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}, \qquad j = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad k = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{5.63}$$

We can check that these matrices satisfies the quaternionic algebra :

$$\begin{aligned} i^2 = j^2 = k^2 &= -1 \\ ij = -ji &= k, \\ jk = -kj &= i, \\ ki = -ik &= j. \end{aligned} \tag{5.64}$$

The algebra $\mathrm{Cl}(2) = \mathbb{H}$ is represented by $\tilde{\rho}$ on $\mathbb{C}^2$ by the **Pauli matrices** $1, i, j, k$ which are given by (5.63).

5.4.5.2 The maps α and τ

What are the matrices which represent V ? These are $\tilde{\rho}(e_1)$ and $\tilde{\rho}(e_2)$. Thus we can write $V = \mathrm{Span}_{\mathbb{R}}\{\gamma_1, \gamma_2\} = \mathrm{Span}_{\mathbb{R}}\{j, k\}$, or

$$V = \left\{ \begin{pmatrix} 0 & \xi \\ -\overline{\xi} & 0 \end{pmatrix} \,:\, \xi \in \mathbb{C} \right\}.$$

As before, α is the unique homomorphic extension to Cl(2) of $-\mathrm{id}$ on V. From the definitions, we get $\alpha(j) = -j$, $\alpha(k) = -k$. The extension present no difficult. For example: $\alpha(i) = \alpha(jk) = \alpha(j)\alpha(k) = jk = i$, but $\alpha(jk) = \alpha(i)$; then $\alpha(i) = i$. The same gives $\alpha(1) = 1$.

The case of τ is treated in similar way. We find: $\tau(j) = j$, $\tau(k) = k$, $\tau(i) = -i$, $\tau(1) = 1$.

Now, we can find the group $\Gamma_{(2)}$. The condition for $x \in \mathrm{Cl}(2)$ to be in $\Gamma_{(2)}$ is $\alpha(x)yx^{-1}$ to belongs to V for all $y \in V$. We put

$$x = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}, \qquad \alpha(x) = \begin{pmatrix} \alpha & -\beta \\ \overline{\beta} & \overline{\alpha} \end{pmatrix}.$$

A typical y in V is

$$y = \begin{pmatrix} 0 & \eta \\ -\overline{\eta} & 0 \end{pmatrix}.$$

A few computation gives:

$$\alpha(x)yx^{-1} = \frac{1}{|\alpha|^2 + |\beta|^2} \begin{pmatrix} \alpha\eta\overline{\beta} + \beta\overline{\eta\alpha} & \alpha\alpha\eta - \beta\beta\overline{\eta} \\ \overline{\beta\beta}\eta - \overline{\alpha\alpha}\eta & \eta\alpha\overline{\beta} + \overline{\alpha\eta}\beta \end{pmatrix}.$$

If we impose it to be of the form $\begin{pmatrix} 0 & \xi \\ -\overline{\xi} & 0 \end{pmatrix}$ for all $\eta \in \mathbb{C}$, we get, for all $\eta \in \mathbb{C}$, $\mathrm{Re}(\overline{\alpha}\beta\overline{\eta}) = 0$, which implies $\overline{\alpha}\beta = 0$. So we conclude:

$$\Gamma_{(2)} = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix}, \begin{pmatrix} 0 & \beta \\ -\overline{\beta} & 0 \end{pmatrix} \ : \ \alpha, \beta \in \mathbb{C} \text{ not both equals zero} \right\}.$$

Be careful on a point: $\Gamma_{(2)}$ is the *multiplicative* group generated by these two matrices, not the additive one.

5.4.5.3 The spin group

It present no difficult to find that

$$\Gamma^+_{(2)} = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix} \ : \ \alpha \neq 0 \right\}. \tag{5.65}$$

The **spin group** is made of elements of $\Gamma^+_{(2)}$ which satisfy $\tau(x) = x^{-1}$. We know that $\tau \begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix} = \begin{pmatrix} \overline{\alpha} & 0 \\ 0 & \alpha \end{pmatrix}$ and that $\begin{pmatrix} \alpha & 0 \\ 0 & \overline{\alpha} \end{pmatrix}^{-1} = \frac{1}{\alpha\overline{\alpha}} \begin{pmatrix} \overline{\alpha} & 0 \\ 0 & \alpha \end{pmatrix}$. Thus the condition $\tau(x) = x^{-1}$ becomes $|\alpha|^2 = 1$. The first conclusion is that

$$\mathrm{Spin}(2) = U(1). \tag{5.66}$$

A typical s in Spin(2) is

$$s = e^{i\theta} = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

The next point is to see the action of Spin(2) on V. The action of $s \in \mathrm{Spin}(2)$ on a vector $v \in V$ is still defined by $s \bullet v = \chi(s)v = \alpha(s) \cdot v \cdot s^{-1}$. More explicitly:

$$\chi(s)v = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \begin{pmatrix} 0 & z \\ -\overline{z} & 0 \end{pmatrix} \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} = \begin{pmatrix} 0 & e^{2i\theta}z \\ -e^{-2i\theta}\overline{z} & 0 \end{pmatrix}, \tag{5.67}$$

where the matrix $\begin{pmatrix} 0 & z \\ \overline{z} & 0 \end{pmatrix}$ denotes the representation of the vector v of V. This equality can be written $e^{i\theta} \cdot v = e^{2i\theta} v$. If we note $v = v_1 + iv_2 = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, we get

$$e^{2i\theta} \bullet v = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Therefore, we can write

$$\chi(e^{i\theta}) = \begin{pmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}.$$

So χ projects $U(1)$ into SO(2) with a kernel $\mathbb{Z}_2$, for this reason, we say that $U(1)$ is a **double covering** of SO(2). We note it

$$\mathbb{Z}_2 \to U(1) \xrightarrow{\chi} \mathrm{SO}(2). \tag{5.68}$$

5.5 Clifford modules

References: [39, 40].

Let M be a manifold. We denote by $\mathrm{Cl}^{\mathbb{C}}(M)$ the bundle whose fibre at $x \in M$ is the complex Clifford algebra of the metric g_x : $\mathrm{Cl}^{\mathbb{C}}(M)_x = \mathrm{Cl}^{\mathbb{C}}(g_x)$. We define the important map

$$\begin{aligned} \gamma\colon \Gamma(M, \mathrm{Cl}^{\mathbb{C}}(M)) &\to \mathfrak{B}(\mathscr{H}) \\ \gamma(dx^\mu) &\mapsto \gamma^\mu(x) \end{aligned} \tag{5.69}$$

which can be extended to the whole Clifford algebra.

Let V be a vector space endowed with a bilinear symmetric form. We consider $\mathrm{Cl}(V)$, the corresponding Clifford algebra. A **Clifford module** is a real vector space E with a $\mathbb{Z}_2$-graduation and a morphism

$$\rho_E\colon\ \mathrm{Cl}(V) \to \mathrm{End}(E)$$

of $\mathbb{Z}_2$-graded vector spaces. It is defined by a linear map $\rho_E\colon V \to \mathrm{End}(V)$ such that

$$\rho_E(v)\rho_E(w) + \rho_E(w)\rho_E(v) = B(v,w)\,\mathrm{id} \tag{5.70}$$

for every v, $w \in E$. The element $\rho_E(x)v$ will often be denoted by $x \cdot v$ and the operation ρ_E is the **Clifford multiplication**. The **dual module** E^* is defined by $\rho_{E^*}(x) = \rho_E(x^t)^*$, i.e.

$$\langle \rho_{E^*}(x)\psi, v\rangle = (-1)^{|\psi||x|}\langle \psi, \rho_E\big(\tau(x)\big)v\rangle \tag{5.71}$$

for every $\psi \in E^*$ and $v \in E$. Here

Let $\mathfrak{A}$ be a $\mathbb{Z}_2$-graded subalgebra of $\mathrm{Cl}(V)$ and E_1, a $\mathfrak{A}$-module. Then the space

$$E = \mathrm{Ind}_{\mathfrak{A}}^{\mathrm{Cl}(V)}(E_1) = \mathrm{Cl}(V) \otimes_{\mathfrak{A}} E_1$$

has a structure of Clifford module, the **induced module**. The tensor product $\otimes_{\mathfrak{A}}$ is the usual one modulo the subspace spanned by elements of the form

$$x \otimes a \cdot y - xa \otimes y$$

for every x, $a \in \mathrm{Cl}(V)$ and $y \in E_1$. In a similar way, if E is a complex vector space we have a notion of $\mathrm{Cl}^{\mathbb{C}}(V)$-module.

Let $x \in \mathrm{Cl}(V)$ be such that $x^2 = 1$. In that case the Clifford multiplication $\rho_E(x)$ decomposes E in eigenspaces

$$E^{\pm} = \frac{1}{2}\big(1 \pm \rho_E(x)\big)E.$$

If V is a n-dimensional vector space with an oriented orthonormal basis $\{e_1, \ldots, e_n\}$, the algebra $\mathrm{Cl}(V)$ has a **volume element** $\omega = e_1 e_2 \ldots e_n$ which does not depend on the choice of the basis. The volume element squares to

$$\omega^2 = (-1)^{n(n+1)/2}. \tag{5.72}$$

In the complex case, we consider the complex vector space $V^{\mathbb{C}}$ and the complex Clifford algebra $\mathrm{Cl}^{\mathbb{C}}(V) = \mathrm{Cl}(V) \otimes_{\mathbb{R}} \mathbb{C}$, and the volume element is defined as

$$\omega_{\mathbb{C}} = i^{[(n+1)/2]}\omega. \tag{5.73}$$

where $[x]$ is denotes the integer part of x. Performing a separate computation for n even or odd, it is easy to see that in both case,

$$\omega_{\mathbb{C}}^2 = 1. \tag{5.74}$$

So in the complex case we always have an element in $\mathrm{Cl}(V)$ which squares to 1, and a $\mathrm{Cl}^{\mathbb{C}}(V)$-module W always accepts a decomposition as $W^{\pm} = \frac{1}{2}(1 + \omega_{\mathbb{C}})W$.

One says that a representation ρ of $\mathrm{Cl}(V)$ on W is **reducible** if there exists a splitting $W = W_1 \oplus W_2$ such that $\rho(\mathrm{Cl}(V))W_i \subset W_i$. If the representation is not reducible, it is said to be irreducible. Two representations $\rho_j\colon\ \mathrm{Cl}(V) \to \mathrm{End}(W_j)$ are **equivalent** if there exists a linear isomorphism $F\colon W_1 \to W_2$ such that $F \circ \rho_1(x) \circ F^{-1} = \rho_2(x)$ for every $x \in \mathrm{Cl}(V)$.

Proposition 5.36.
The real Clifford algebra has

$$\begin{cases} 2 & \text{if } n+1 = 0 \mod 4 \\ 1 & \text{otherwise} \end{cases}$$

inequivalent irreducible representations. The complex Clifford algebra $\mathrm{Cl}^{\mathbb{C}}(V)$ *has*

$$\begin{cases} 2 & \text{if } n \text{ is odd} \\ 1 & \text{if } n \text{ is even} \end{cases}$$

inequivalent irreducible representations.

Proof. No proof. □

If M is a manifold, we denote by $\mathrm{Cl}(M) = \mathrm{Cl}(TM)$ the bundle whose fiber at x is the Clifford algebras of T_xM. We consider an orthonormal basis $\{e_i\}$ and if Σ is a multi-index $\{1 \leqslant \sigma_1, \ldots, \leqslant \sigma_t \leqslant m\}$, we pose $e_\Sigma = e_{\sigma_1} \ldots e_{\sigma_t} \in \mathrm{Cl}(M)$. By convention, $e_\varnothing = 1$. Since the elements e_i are ordered, they provide an orientation:

$$d\,\mathrm{Vol} = e_1 \wedge \ldots \wedge e_m \in \bigwedge^m(M). \tag{5.75}$$

Since the map $e_{\sigma_1} \wedge \ldots \wedge e_{\sigma_t} \mapsto e_{\sigma_1 \ldots e_{\sigma_t}}$ is an isomorphism between $\mathrm{Cl}(M)$ and $\bigwedge(M)$, we say that $d\,\mathrm{Vol} \in \mathrm{Cl}(M)$. Now we define

$$\kappa = i^{-[(m+1)/2]} d\,\mathrm{Vol},$$

which is nothing else that the volume form normalised in such a way that $\kappa^2 = 1$. If m is even, it anti-commutes with TM, and if m is odd, it commutes with TM.

Let V be a m-dimensional real vector space, and $\mathrm{Cl}^{\mathbb{C}}(V)$, the corresponding complex Clifford algebra.

Lemma 5.37.
Every $\mathrm{Cl}^{\mathbb{C}}(V)$*-module accepts an unique decomposition as sum of irreducible representations as follows*

(i) if $m = 2n$*, there exists one and only one irreducible* $\mathrm{Cl}^{\mathbb{C}}(V)$*-module* Δ *and* $\dim(\Delta) = 2n$,

(ii) if $m = 2n+1$*, we have two inequivalent irreducible modules* $\Delta_\pm$ *with* $\gamma(\kappa) = \pm 1$ *on* $\Delta_\pm$ *and* $\dim(\Delta_\pm) = 2^n$.

Proof. No proof. □

Let V be a vector bundle over M. A structure of $\mathrm{Cl}(M)$-module over V is a morphism of unital algebra $\gamma\colon \mathrm{Cl}(M) \to \mathrm{End}(V)$. When one has a basis $\{e_i\}$ of V, we pose $\gamma_i = \gamma(e_i)$. The following lemma is the lemma 1.2 of [40].

Lemma 5.38.
Let V *be a* $\mathrm{Cl}(V)$*-module and* $\{e_i\}$*, an orthonormal basis for* TM *on a contractible open set* V*. Then there exists a local frame for* V *such that the matrices* $\gamma(e_i)$ *are constant.*

We also define $\gamma^i = \gamma(dx^i) = g^{ij}\gamma_j$. One easily proves that

$$\gamma^i\gamma^j + \gamma^j\gamma^i = -2g^{ij} \tag{5.76}$$

where (g^{ij}) is the inverse matrix of (g_{ij}). If the endomorphisms γ_i are constant in the basis $\{e_i\}$, then the endomorphisms γ^i are constant in the basis $\{f_i = g_{ki}e_k\}$.

5.6 Spin structure

We consider a (pseudo-)Riemannian manifold (M, g) with metric signature (p, q), and $\mathrm{SO}(M)$, its frame bundle; it admits a $\mathrm{SO}(p, q)$-principal fibre bundle structure which is well defined by the metric g (see 4.4.4.1).

Definition 5.39.
We say that (M, g) *is a* ***spin manifold*** *if there exists a* $\mathrm{Spin}(p, q)$*-principal bundle* P *over* M *and a principal bundle homomorphism* $\varphi\colon P \to \mathrm{SO}(M)$ *which induced covering for the structure groups is* χ*, i.e.* $\varphi(\xi \cdot s) = \varphi(\xi) \cdot \chi(s)$*. A choice of* P *and* φ *is a* ***spin structure*** *on* M*.*

$$\begin{tikzcd}
\mathrm{Spin}(p,q) \arrow[r, rightsquigarrow] & P \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & \mathrm{SO}(M) \arrow[dl, "p"] & \mathrm{SO}(p,q) \arrow[l, rightsquigarrow] \\
& & M & &
\end{tikzcd}$$

The wavy arrows mean “structural group of”.

Remark 5.40.
When we will use the concept of spin structure in the physical oriented chapters, we will naturally use $\mathrm{SL}(2,\mathbb{C})$ *as group instead of* $\mathrm{Spin}(p,q)$*. The isomorphism* $\mathrm{SL}(2,\mathbb{C}) \simeq \mathrm{Spin}(1,3)$ *is proved in [37]. A physical motivation of such a structure is given at page 257.*

5.6.1 Example: spin structure on the sphere S^2

It is no difficult to see that $\mathrm{SO}(S^2) \simeq \mathrm{SO}(3)$. Indeed, each element of $\mathrm{SO}(S^2)$ is described by three orthonormal vectors: one which point to an element x of S^2 and two which gives a basis of T_xS^2. The action $\mathrm{SO}(3) \times S^2 \to S^2$ is transitive, and the stabilizer of any element is $\mathrm{SO}(2)$.

We define $\alpha\colon \mathrm{SO}(3)/\mathrm{SO}(2) \to S^2$ by $\alpha(g\,\mathrm{SO}(2)) = g$. Proposition **??** shows that α is a diffeomorphism. Then

$$S^2 = \frac{\mathrm{SO}(3)}{\mathrm{SO}(2)}.$$

On the other hand, we know that

$$T_eSU(2) = su(2) = \left\{ \begin{pmatrix} ix & \xi \\ -\bar{\xi} & -ix \end{pmatrix} : \xi \in \mathbb{C}, x \in \mathbb{R} \right\}. \tag{5.77}$$

It is a classical result that $\mathfrak{su}(2) \simeq \mathbb{R}^3$ not only as set but also as metric space with the identification

$$\langle X, Y \rangle = -\frac{1}{2}\operatorname{Tr}(XY),$$

for all X, $Y \in su(2)$. As we are in matrix groups, we know (see [10] to get more details) that $Ad_xY = xYx^{-1}$. In our case, this gives the formula

$$\langle Ad(g)X, Ad(g)Y \rangle = \langle X, Y \rangle.$$

We can now state the result for S^2.

Proposition 5.41.
The manifold S^2 *with the usual metric induced from* $\mathbb{R}^3$ *admits the following spin structure:*

$$\begin{tikzcd} \mathrm{Spin}(2) \arrow[r, squiggly] & SU(2) \arrow[rr, "\varphi = Ad"] \arrow[dr, "\pi", "U(1)"'] & & \mathrm{SO}(3) \arrow[dl, "p"', "\mathrm{SO}(2)"] \\ & & S^2 & \end{tikzcd} , \tag{5.78}$$

where the arrow $X \xrightarrow[G]{f} Y$ *means that* G *is the kernel of the map* $f\colon X \to Y$.

Proof. First, let us precise the concept of frame bundle for S^2, and how it is well described by $\mathrm{SO}(3)$. Let $\{e_1, e_2, e_3\}$ be the canonical basis of $\mathbb{R}^3$. To $A \in \mathrm{SO}(3)$, we make correspond the basis $\{Ae_2, Ae_3\}$ at the point Ae_1 of S^2. The projection $p\colon \mathrm{SO}(3) \to S^2$ is then defined by $p(A) = Ae_1$. It is clear that we will define the map $\pi\colon SU(2) \to S^2$ in the same way: $\pi(U) = p(Ad(U))$.

For the rest of the demonstration, we will use the "$su(2)$ description" of $\mathbb{R}^3$ given by (5.77) with $\xi = y + iz$.

Now, let us show that $\pi\colon SU(2) \to S^2$ is a $\mathrm{Spin}(2)$-principal bundle. Since we had already shown that $\mathrm{Spin}(2) \simeq U(1)$, we define the right action of $\mathrm{Spin}(2)$ on $SU(2)$ by right multiplication: $U \cdot s = Us$ with $s = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$. It is clear that $\pi(Us) = \pi(U)$:

$$Ad(Us)e_1 = (Us)\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} s^{-1}U^{-1} = Us\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} s^{-1}U^{-1}, \tag{5.79}$$

because $\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ is the vector e_1 in the "$su(2)$ description" of $\mathbb{R}^3$.

In order for $\pi\colon SU(2) \to S^2$ to be a $\mathrm{Spin}(2)$-principal bundle, we still need to show that for all $x \in S^2$,

$$\pi^{-1}(x) = \left\{ \xi \cdot g \text{ st } g \in \mathrm{Spin}(2)\, \forall \xi \in \pi^{-1}(x) \right\}.$$

Take A, $B \in \pi^{-1}(x)$, i.e. $Ae_1 = Be_1 = x$. We need to find a $s \in \mathrm{Spin}(2)$ such that

$$A = B \cdot s. \tag{5.80}$$

The matrices A and B are such that

$$B^{-1}A\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}A^{-1}B = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \tag{5.81}$$

This implies that $B^{-1}A \in \mathrm{Spin}(2)$. As Ad is surjective from $SU(2)$ into $\mathrm{SO}(3)$, a general C in $\mathrm{SO}(3)$ which acts on e_1 can be written Ue_1U^{-1} for $U \in SU(2)$ such that $Ad(U) = C$. The condition (5.81) becomes

$$\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} \overline{\alpha} & -\beta \\ \overline{\beta} & \alpha \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

which implies $\alpha = e^{i\theta}$, $\beta = 0$. Then $B^{-1}A$ belongs to $\mathrm{Spin}(2)$, and $s = B^{-1}A$ fulfills the condition (5.80).

What about the induced covering for the structural groups ? The structural group of $\pi\colon SU(2) \to S^2$ is $\mathrm{Spin}(2)$, while the one of $p\colon \mathrm{SO}(3) \to S^2$ is $\mathrm{SO}(2)$. Indeed, for each $x \in S^2$, $\mathrm{SO}(2)$ acts on T_xS^2, leaving x unchanged. We have the following associations:

$$U \in SU(2) \xrightarrow{\varphi} A \in \mathrm{SO}(3),$$

the matrix A being defined by $A \cdot X = UXU^{-1}$. For $s \in \mathrm{Spin}(2)$ we of course also have

$$Us \in SU(2) \xrightarrow{\varphi} As \in \mathrm{SO}(3),$$

with $As \cdot X = UsXs^{-1}U^{-1}$. As we act by $\mathrm{Spin}(2)$ on $SU(2)$, in the fibres of $\mathrm{SO}(3)$, the action of $\mathrm{Spin}(2)$ is –via φ– the composition with $X \to sXs^{-1}$. But this is exactly $\chi(s)X$ because $\alpha(s) = s$, since $s \in \mathrm{Spin}(2)$. $\square$

5.6.2 Spinor bundle

Let us take once again the spin structure on the (pseudo-)Riemannian manifold (M, g):

$$\begin{tikzcd}
\mathrm{Spin}(p,q) \arrow[r, rightsquigarrow] & P \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & \mathrm{SO}(M) \arrow[dl, "p"] & \mathrm{SO}(p,q) \arrow[l, leftsquigarrow] \\
& & M & &
\end{tikzcd}$$

with $\varphi(\xi \cdot g) = \varphi(\xi) \cdot \chi(g)$.

Let us define $S = \Lambda W$, and $\mathcal{S} = P \times_\rho S$. Take $\rho\colon \mathrm{Spin}(p,q) \times \mathcal{S} \to \mathcal{S}$, $\rho(g, s) = \tilde{\rho}(g)s$, where $\tilde{\rho}$ is the spinor representation of $\mathrm{Spin}(p,q)$ on S. We also have $\chi\colon \mathrm{Spin}(p,q) \to \mathrm{SO}_0(p,q)$, $\chi(g)v = \alpha(g) \cdot v \cdot g^{-1}$, with $\alpha(g) = g$ for $g \in \mathrm{Spin}(p,q)$.

The **spinor bundle** is the associated bundle

$$\mathcal{S} = P \times_\rho S \to M \tag{5.82}$$

A **spinor field** is an element of $\Gamma(\mathcal{S})$, the space of section of the spinor bundle.

On $\mathrm{SO}(M)$, we look at a connection 1-form $\alpha \in \Omega^1(\mathrm{SO}(M), so(\mathbb{R}^m))$, and, if $T(M)$ is the tensor bundle over M, we define a covariant derivative $\nabla^\alpha\colon \mathfrak{X}(M) \times T(M) \to T(M)$ by

$$\widehat{\nabla^\alpha_X s} = \overline{X}\hat{s},$$

for any $s \in T(M)$. See theorem 4.27, and the fact that $T(M)$ can be see as an associated bundle; it is explicitly done for $\mathfrak{X}(M)$ at page 157.

As seen in point 4.11.2.2, an automatic property of this connection is $\nabla^\alpha g = 0$ if g is the metric of M. The **Levi-Civita connection** is the unique[3] such connection which is torsion-free: $T^{\nabla^\alpha} = 0$.

Proposition 5.42.
The 1-form $\tilde{\alpha} = \varphi^*\alpha \in \Omega^1(P, so(\mathbb{R}^m))$ defines a connection on P. See definition 4.21 and theorem 4.27.

Proof. Let us denote by R_g the right action of $g \in \mathrm{Spin}(p,q)$ on P (*id est* $R_g\xi = \xi \cdot g$), and by $R_u^{\mathrm{SO}(M)}$ the right action of $u \in \mathrm{SO}(p,q)$ on $\mathrm{SO}(M)$. We have to check the usual two conditions of a connection.
First condition. The first one is:

$$(R_g^*\tilde{\alpha})_\xi(\Sigma) = Ad(g^{-1})(\tilde{\alpha}_\xi(\Sigma)),$$

[3] We will not prove unicity.

for all $\xi \in P$, and $\Sigma \in T_\xi P$. In order to check this, we first remark that $\varphi \circ R_g = R^{\mathrm{SO}(M)}_{\chi(g)} \circ \varphi$. Indeed, for all $\xi \in P$, definition 5.39 gives us $\varphi(R_g\xi) = \varphi(\xi \cdot g) = \varphi(\xi) \cdot \chi(g)$. With this, we can make the following computation:

$$\begin{aligned} R_g^*\tilde{\alpha} &= R_g^*\varphi^*\alpha = (\varphi \circ R_g)^*\alpha = (R^{\mathrm{SO}(M)}_{\chi(g)} \circ \varphi)^*\alpha \\ &= \varphi^* R^{\mathrm{SO}(M)*}_{\chi(g)}\alpha = \varphi^*(Ad(\chi(g)^{-1}) \circ \alpha). \end{aligned} \tag{5.83}$$

The last equality comes from the fact that α is a connection 1-form. As we are in matrix groups, we have $Ad(g)x = gxg^{-1}$, so

$$[Ad(\chi(g))x]v = [\chi(g)x\chi(g)^{-1}]v = \chi(g)[xg^{-1}vg] = gxg^{-1}. \tag{5.84}$$

In the first line, the product is the usual matrix product which can be seen as operator composition.

But $(Ad(g)x)v = gxg^{-1}v$. Then $Ad(g) = Ad(\chi(g))$, if we identify $\mathfrak{spin}(p,q) \simeq \mathfrak{so}(p,q)$ by proposition 5.34. Moreover, the action of Ad is linear, so it commutes with φ^*. With these remarks, we can continue the computation (5.83):

$$\varphi^*(Ad(\chi(g)^{-1}) \circ \alpha) = \varphi^*(Ad(g^{-1}) \circ \alpha) = Ad(g^{-1}) \circ \varphi^*\alpha = Ad(g^{-1}) \circ \tilde{\alpha}. \tag{5.85}$$

This proves the first condition.

Second condition. The second one is $\tilde{\alpha}(A^*_\xi) = -A$ with the definition (4.100). This is also a computation. First remark

$$\tilde{\alpha}_\xi(A^*_\xi) = (\varphi^*\alpha)_\xi(A^*_\xi) = \alpha_{\varphi(\xi)}(\varphi_{*\xi}A^*_\xi).$$

We compute $\varphi_{*\xi}A^*$ with lemma 2.13:

$$\begin{aligned} \varphi_{*\xi}A^* &= \frac{d}{dt}\varphi(\xi \cdot \exp -tA)\Big|_{t=0} = \frac{d}{dt}(R^{\mathrm{SO}(M)}_{\chi(\exp -tA)} \circ \varphi)(\xi)\Big|_{t=0} \\ &= \frac{d}{dt}\varphi(\xi) \cdot \chi(\exp -tA)\Big|_{t=0} = \frac{d}{dt}\varphi(\xi) \cdot \exp(-td\chi_e A)\Big|_{t=0} = (d\chi_e A)^*_{\varphi(\xi)}. \end{aligned} \tag{5.86}$$

But $d\chi_e = \mathrm{id}_{so(p,q)}$, thus $\varphi_{*\xi}A^* = A^*_{\varphi(\xi)}$. The whole makes that:

$$\tilde{\alpha}_\xi(A^*_\xi) = \alpha_{\varphi(\xi)}(\varphi_{*\xi}A^*_\xi) = \alpha_{\varphi(\xi)}(A^*_{\varphi(\xi)}) = -A.$$

This completes the proof. □

Definition 5.43.
*This connection 1-form on P is called the **spinor connection**. It gives us a covariant derivative on any associated bundle and in particular on the spinor bundle, $\tilde{\nabla}\colon \mathfrak{X}(M) \times \Gamma(\mathcal{S}) \to \Gamma(\mathcal{S})$.*

Proposition 5.44.
If $X, Y \in \mathfrak{X}(M)$ are such that $X_x = Y_x$, then for all $s \in \Gamma(\mathcal{S})$,

$$(\tilde{\nabla}_X s)(x) = (\tilde{\nabla}_Y s)(x).$$

Proof. We just have to show that for all vector field Z such that $Z_x = 0$, $(\tilde{\nabla}_Z s)(x) = 0$. Such a Z can be written as $Z = fZ'$ for a function f on M which satisfies $f(x) = 0$. We have:

$$\tilde{\nabla}_Z s = \tilde{\nabla}_{fZ'}s = f\tilde{\nabla}_{Z'}s,$$

which is obviously zero at x. □

Let $x \in M$ and $\{e_{\alpha\, x}\}$ be an orthonormal basis of T_xM. We can extend it to $\{e_\alpha\}$, a local basis field around x such that e_α is a section of the frame bundle (in other words, we ask the extension to be smooth). The claim of proposition 5.44 is that $\tilde{\nabla}_{e_\alpha}(x)$ is an element of $\mathcal{S}_x$ which doesn't depend on the extension.

5.7 Dirac operator ☺

5.7.1 Preliminary definition

Let M be a m-dimensional (pseudo)Riemannian manifold with its spin structure

$$\begin{tikzcd} \mathrm{Spin}(p,q) \arrow[r, squiggly] & P \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & \mathrm{SO}(M) \arrow[dl, "p"] & \mathrm{SO}(p,q) \arrow[l, squiggly] \\ & & M & & \end{tikzcd}$$

where φ satisfies $\varphi(\xi \cdot g) = \varphi(\xi) \cdot \chi(g)$.

Recall that for any vector space, one can see $\operatorname{End} V = V^* \otimes V$ with the definition $(v^* \otimes v)w = (v^* w)v$. This allows us to define an action of $\mathrm{Spin}(p, q)$ on $\operatorname{End} S$ by defining an action of $\mathrm{Spin}(p, q)$ on S and S^* separately. We know the action

$$\begin{aligned} \mathrm{Spin}(p, q) \times S &\to S \\ (g, v) &\mapsto \tilde{\rho}(g)v, \end{aligned} \tag{5.87}$$

and as action on S^*, we take the dual one

$$\begin{aligned} \mathrm{Spin}(p, q) \times S^* &\to S^* \\ g \cdot \alpha &= \alpha \circ \tilde{\rho}(g^{-1}) \end{aligned} \tag{5.88}$$

for all $g \in \mathrm{Spin}(p, q)$ and $\alpha \in S^*$.

Now we can make the following computation with $g \in \mathrm{Spin}(p, q)$, $\alpha \in S^*$ and $v \in S$, using the fact that $\tilde{\rho}$ is linear:

$$\begin{aligned} [g \cdot (\alpha \otimes v)]w &= [(\alpha \circ \tilde{\rho}(g^{-1}))w]\tilde{\rho}(g)v \\ &= \tilde{\rho}\left([(\alpha \circ \tilde{\rho}(g^{-1}))w]g\right) v \\ &= \left[\tilde{\rho}(g) \circ (\alpha \otimes v) \circ \tilde{\rho}(g^{-1})\right]w. \end{aligned} \tag{5.89}$$

Then we write the action of $\mathrm{Spin}(p, q)$ on $\operatorname{End} S$ by ($A \in \operatorname{End} S$)

$$g \cdot A = \tilde{\rho}(g) \circ A \circ \tilde{\rho}(g^{-1}). \tag{5.90}$$

Notice that this definition is the one required in condition (4.50).

The tangent bundle T_xM is given with a metric g_x. As usual, we build $S_x = \Lambda W_x$, a completely isotropic subspace of T_xM with respect to the metric g_x, and a representation

$$\tilde{\rho}_x \colon T_xM \to \operatorname{End}(\Lambda W_x)$$

The first step in the definition of $\gamma(X)$ is to build $\hat{a}_X \colon P \to \operatorname{End}(\Lambda W)$ setting[4] $\hat{a}_X(p) = \tilde{\rho}(\hat{X}_{\varphi(p)})$.

Lemma 5.45.
The function $\hat{a}$ is equivariant, i.e. it satisfies

$$\hat{a}_X(p \cdot g) = g^{-1} \cdot \hat{a}_X(p) \tag{5.91}$$

for all $g \in \mathrm{Spin}(p, q)$.

Proof. It is no more than a simple computation using the equivariance of $\hat{X}$. Indeed:

$$\begin{aligned} \hat{a}_X(p \cdot g) &= \tilde{\rho}(\hat{X}_{\varphi(p \cdot g)}) = \tilde{\rho}(\hat{X}_{\varphi(p)\chi(g)}) = \tilde{\rho}(\chi(g^{-1}) \cdot \hat{X}_{\varphi(p)}) \\ &= \tilde{\rho}(g^{-1} \cdot \hat{X}_{\varphi(p)} \cdot g) = \tilde{\rho}(g^{-1}) \circ \tilde{\rho}(\hat{X}_{\varphi(p)}) \circ \tilde{\rho}(g) \\ &= g^{-1} \cdot \hat{a}_X(p). \end{aligned} \tag{5.92}$$

In the fourth line, the dots mean the Clifford product, and the last equality comes from the definition of the action (5.90) of $\mathrm{Spin}(p, q)$ on $\operatorname{End} S$. □

From the discussion of section 4.5.2, the function $\hat{a}_X \colon P \to \operatorname{End} S$ defines a section $a_X \colon M \to \operatorname{End} \mathcal{S}$. We define $\gamma \colon \mathfrak{X}(M) \to \operatorname{End} \Gamma(\mathcal{S})$ by

$$\gamma(X) = a_X. \tag{5.93}$$

We immediately have

$$\widehat{\gamma(X)}(p) = \tilde{\rho}(\hat{X}_{\varphi(p)})$$

for any $p \in P$. If we define

$$\widehat{\gamma \cdot a_X}(p) = \widehat{\gamma(X)}(p), \tag{5.94}$$

the map γ can be seen as an action on the section of $\mathcal{S}$. Indeed, $\widehat{\gamma \cdot s_X}$ is an equivariant function:

$$\begin{aligned} \hat{\gamma}(p \cdot g)(\hat{a}_X(p \cdot g)) &= \rho(g)^{-1}\hat{\gamma}(p)\rho(g)\rho(g^{-1})\hat{a}_X(p) \\ &= \rho(g)^{-1}\hat{\gamma}(p)\hat{a}_X(p) \\ &= \rho(g^{-1})\widehat{\gamma \cdot a_X}(p), \end{aligned} \tag{5.95}$$

so that

$$\widehat{\gamma \cdot a_X}(p) = \rho(g^{-1})\widehat{\gamma \cdot a_X}(p).$$

The map $\widehat{\gamma \cdot a_X} \colon P \to \operatorname{End} \Lambda W$ defined by (5.94) is equivariant, and thus defines a section $\gamma \cdot a_X \in \Gamma(\mathcal{S})$, as seen in the section 4.5.2.

[4]See subsection 4.5.4 for the definition of $\hat{X}$.

5.7.2 Definition of Dirac

If we consider a basis $\{e_\alpha\}$ of TM, *i.e.* m sections $e_\alpha\colon M \to TM$ such that for all x in M, the set $\{e_{\alpha x}\}$ is a basis of T_xM, we note $\gamma^\alpha := \gamma(e_\alpha) \in \operatorname{End}(\mathcal{S})$.

Remark 5.46.
This is not always globally possible. The example of the sphere is given in subsection **??**.

For any $s \in \Gamma(\mathcal{S})$, we consider the local[5] section ψ of $\mathcal{S}$ given by

$$\psi(x) = \sum_{\alpha\beta} g_x(e_\alpha, e_\beta)\gamma_x^\beta(\tilde{\nabla}_{e_\alpha} s)(x).$$

For each $x \in M$, take a A_x in[6] $SO(g_x)$, and consider the new basis $e'_\alpha = A_\alpha{}^\beta e_\beta$. As A is an isometry, $g_x(e'_\alpha, e'_\beta) = g_x(e_\alpha, e_\beta)$; and since $\tilde{\rho}$ is linear, $\gamma_x'^\alpha = \tilde{\rho}_x(e'_{\alpha x}) = A_\alpha{}^\beta \tilde{\rho}(e_{\beta x}) = A_\alpha{}^\beta \gamma_x^\beta$. In the new basis, the section reads:

$$\begin{aligned}
\psi(x) &= \sum_{\alpha\beta\eta\sigma} g_x(e_\alpha, e_\beta) A_\beta{}^\sigma \gamma_x^\sigma (\tilde{\nabla}_{A_\alpha{}^\eta e_\eta} s)(x) \\
&= \sum_{\alpha\beta\eta\sigma} (A^t)^\eta{}_\alpha g_{\alpha\beta}(x) A_\beta{}^\sigma \gamma_x^\sigma (\tilde{\nabla}_{e_\eta} s)(x) \\
&= \sum_{\eta\sigma} g_x(e_\eta, e_\sigma)\gamma_x^\sigma(\tilde{\nabla}_{e_\eta} s)(x),
\end{aligned} \tag{5.96}$$

where we used the fact that $A^t g A = g$ and that all the $A_\alpha{}^\beta$ are C^∞ functions on M, so that $\tilde{\nabla}_{A_\alpha{}^\beta X} = A_\alpha{}^\beta \tilde{\nabla}_X$. This shows that $\psi(x)$ doesn't depend on the choice of the basis, so it defines a section from the data of s alone.

The **Dirac operator** $\mathcal{D}\colon \Gamma(\mathcal{S}) \to \Gamma(\mathcal{S})$ acting on a spinor field is defined by

$$(\mathcal{D}s)(x) = g_x(e_\alpha, e_\beta)\gamma_x^\beta(\tilde{\nabla}_{e_\alpha} s)(x). \tag{5.97}$$

Proposition 5.47.
If the field of basis $e_\alpha \in \mathfrak{X}(M)$ is everywhere an orthonormal basis, the Dirac operator reads

$$(\mathcal{D}s)(x) = g_{\alpha\beta}\gamma^\alpha(\tilde{\nabla}_{e_\beta} s)(x) \tag{5.98}$$

where γ^α is a constant numeric matrix acting on ΛW.

Proof. The building of the Dirac operator begins by considering the vector space T_xM endowed with the metric g_x; then the spinor representation $\tilde{\rho}_x\colon T_xM \to \operatorname{End}(\Lambda W_x)$ where ΛW_x is build from isotropic vectors of T_xM is defined. If the vector fields $e_\alpha \in \mathfrak{X}(M)$ are everywhere orthonormal for the metric g, then we have the matricial equality

$$\tilde{\rho}_x\big((e_\alpha)_x\big)_{ij} = \tilde{\rho}(v_\alpha)_{ij} \tag{5.99}$$

where the left hand side describe the matrix component of a linear operator acting on ΛW_x while in the right hand side we have the matrix component of a linear operator acting on ΛW and v_α is a basis on $\mathbb{R}^n$ with respect to which the metric is the same as the metric g_x in the basis $(e_\alpha)_x$. Let $\hat{\psi}\colon P \to \Lambda W$ be an equivariant function; from definition (5.93) of γ we have

$$\big(\gamma(e_\alpha\hat{\psi})\big)(\xi) = (a_\alpha\hat{\psi})(\xi)$$

where $a_\alpha(\xi) = \tilde{\rho}\Big(\tilde{e}_\alpha\big(\phi(\xi)\big)\Big)$. In this expression, $\tilde{e}_\alpha$ is the equivariant function associated with the vector field $e_\alpha \in \mathfrak{X}(M)$. It is defined in subsection 4.5.4 as

$$\begin{aligned}
\tilde{e}_\alpha\colon SO(M) &\to \mathbb{R}^m \\
b &\mapsto b^{-1}\big((e_\alpha)_{\pi(b)}\big).
\end{aligned} \tag{5.100}$$

So we have $\hat{a}_\alpha\colon P \to \operatorname{End}(\Lambda W)$ defined by

$$\hat{a}_\alpha(\xi) = \tilde{\rho}\big(\varphi(\xi)^{-1} e_\alpha(x)\big)$$

[5] Extensions of e_α do not always globally exist, see remark 5.46.

[6] By $SO(g_x)$, we mean the set of all the matrix A such that $A^t g_x A = g$; A_x is an isometry of (T_xM, g_x). In other words, we consider A as a section of what we could call the "isometry bundle".

with $x = \pi(\xi)$. Now if ξ is any element of $\pi^{-1}(x)$, we have

$$\big(\gamma(e_\alpha)\psi\big)(x) = (a_\alpha\psi)(X) = \big[\xi, \hat{a}_\alpha(\xi)\hat{\psi}(\xi)\big] = \big[\xi, \tilde{\rho}\big(\varphi(\xi)^{-1}e_\alpha(x)\big)\hat{\psi}(\xi)\big].$$

There exists a $g \in \mathrm{Spin}(p,q)$ such that $\varphi(\xi \cdot g) = \mathbb{1}$; taking this element and using equivariance of the latter expression,

$$\big(\gamma(e_\alpha)\psi\big)(x) = \big[\xi \cdot g, \tilde{\rho}\big(e_\alpha(x)\big)\hat{\psi}(\xi \cdot g)\big] = \big[\xi \cdot g, \gamma^\alpha\hat{\psi}(\xi)\big] = [\xi, \gamma^\alpha\hat{\psi}(\xi)]. \tag{5.101}$$

What we proved is that $\big(\gamma e_\alpha\psi\big)(x) = \gamma^\alpha\psi(x)$ is the sense that

$$\widehat{\gamma(e_\alpha)\psi} = \gamma^\alpha\hat{\psi}. \tag{5.102}$$

Hence the Dirac operator reads

$$(\mathcal{D}s)(x) = g_{\alpha\beta}\gamma^\alpha\big(\tilde{\nabla}_{e_\beta}s\big)(x)$$

in the sense that

$$\widehat{\mathcal{D}s} = g_{\alpha\beta}\gamma^\alpha \widehat{\tilde{\nabla}_{e_\beta}s}. \tag{5.103}$$

□

An often more convenient way to write the Dirac operator is to consider an orthonormal basis (so that the metric g and the matrices γ are constant) and to consider the equivariant functions:

$$\widehat{\mathcal{D}\psi} = g_{\alpha\beta}\gamma^\alpha\widehat{\nabla_{e_\alpha}\psi}.$$

This formulation is typically used when one search for Dirac operator on Lie groups. In this case, we choose left invariant vector fields generated by an orthonormal basis of the Lie algebra. The resulting field of basis is everywhere Killing-orthonormal.

Acting on a function $f\colon M \to \mathbb{R}$, it is defined by $\mathcal{D}\colon C^\infty(M) \to C^\infty(M)$,

$$(\mathcal{D}f)(x) = g_x(e_\alpha, e_\beta)\gamma_x^\beta(e_{\alpha x} \cdot f). \tag{5.104}$$

With these definitions, one has

$$(\mathcal{D}(fs))(x) = (f\mathcal{D}s)(x) + (\mathcal{D}f)(x).$$

Indeed,

$$\begin{aligned}
(\mathcal{D}(fs))(x) &= g_{\alpha\beta}\gamma_x^\beta(\tilde{\nabla}_{e_\alpha}fs)(s)\\
&= g_{\alpha\beta}\Big((e_\alpha \cdot f)s(x) + f(x)(\tilde{\nabla}_{e_\alpha}s)(x)\Big)\\
&= f(x)(\mathcal{D}s)(x) + g_{\alpha\beta}\gamma_x^\beta(e_{\alpha x} \cdot f)\\
&= (f\mathcal{D}s)(x) + (\mathcal{D}f)(x).
\end{aligned} \tag{5.105}$$

With that definition, the Dirac operator becomes a derivation of the spinor bundle.

5.8 Example: Dirac operator on $\mathbb{R}^2$ with the euclidian metric

Since the frame bundle $B(M)$ is a principal bundle (see subsection 4.4.4), one can consider some associated bundles on it. We are now going to see that the one given by the definition representation $\rho\colon GL(n,\mathbb{R}) \to GL(n,\mathbb{R})$ on $\mathbb{R}^n$ is the tangent bundle. So we study $B(M) \times_\rho \mathbb{R}^n$. By choosing a basis on each point of M, we identify each T_xM to $\mathbb{R}^n$. An element of $B(M) \times \mathbb{R}^n$ is a pair (b,v) with $b = (\mathbf{b}_1, \ldots, \mathbf{b}_n)$ and $v = (v^1, \ldots, v^n)$. We can identify v to the element of T_xM given by $v = v^i\mathbf{b}_i$.

In order to build the associated bundle, we make the identifications

$$(b,v) \cdot g \sim (b \cdot g, g^{-1}v).$$

Here, by gv we mean the vector whose components are given by $(gv)^i = v^j {g_j}^i$. The tangent vector given by $(b \cdot g, g^{-1}v)$ is $(g^{-1}v)^i(b \cdot g)_i = v^j {(g^{-1})_j}^i {g_i}^k\mathbf{b}_k = v^k\mathbf{b}_k$ So the identification map $\psi\colon B(M) \times_\rho \mathbb{R}^n \to TM$ given by

$$\psi([b,v]) = v^i\mathbf{b}_i$$

is well defined.

The following step is to consider the following spin structure:

$$\mathrm{Spin}(2) \rightsquigarrow \mathbb{R}^2 \times \mathrm{SO}(2) \xrightarrow{\varphi} \mathrm{SO}(\mathbb{R}^2) \leftsquigarrow \mathrm{SO}(2)$$

We have to define the two actions and φ. One of the main result of section 5.4.5 is that $\chi\colon \mathrm{Spin}(2) = U(1) \to \mathrm{SO}(2)$ is surjective. So, we can define the action of Spin(2) on P by

$$(x, b) \cdot s = (x, \chi(s)^{-1} b).$$

On the other hand, an element A in $\mathrm{SO}(\mathbb{R}^2)$ can be written as $A = \{ae_i\}_x$ where e_i is the canonical basis of $T_x\mathbb{R}^2$, and a is a matrix of SO(2). See subsection 4.4.4. For $g \in \mathrm{SO}(2)$, we define

$$A \cdot g = \{g^{-1} a e_i\}_x. \tag{5.106}$$

and $\varphi\colon \mathbb{R}^2 \times \mathrm{SO}(2) \to \mathrm{SO}(\mathbb{R}^2)$ by

$$\varphi(x, b) = \{be_i\}_x.$$

The following shows that these definitions give a spin structure:

$$\varphi((x, b) \cdot s) = \varphi(x, \chi(s)^{-1} b) = \{\chi(s)^{-1} b e_i\}_x = \{be_i\}_x \cdot \chi(s) = \varphi(x, b) \cdot \chi(s). \tag{5.107}$$

5.8.1 Connection on $\mathrm{SO}(\mathbb{R}^2)$

We are searching for a torsion-free connection on the simplest metric space: the euclidian $\mathbb{R}^2$. Thus we will try the simplest choice of horizontal space: we want an horizontal vector to be tangent to a curve of the form $X(t) = \{be_i\}_{x(t)}$. For this reason, we want to define the connection 1-form by $\omega(X) = b'(0)$. For technical reasons which will soon be apparent, we will not exactly proceed in this manner. For $X(t) = \{be_i\}_{x(t)}$, we define

$$\omega(X) = -(b(t)b(0)^{-1})'(0). \tag{5.108}$$

We of course have $\omega(X) = 0$ if and only if $b'(0) = 0$: this choice of ω follows our first idea. In order for ω to be a connection form, we have to verify the two conditions of definition 4.21.

Proposition 5.48.
The 1-form defined by

$$\omega(X) = -(b(t)b(0)^{-1})'(0)$$

for $X = \frac{d}{dt}\{b(t)e_i\}_{x(t)}\Big|_{t=0}$ *is a connection 1-form.*

Proof. Let $A \in \mathrm{SO}(2)$. If $u = \{be_i\}_x$, equation (5.106) gives:

$$A^*_u = \frac{d}{dt}\{e^{-tA} b e_i\}_x\Big|_{t=0},$$

so that $\omega(A^*_u) = -(e^{-tA} b b^{-1})'(0) = A$. This checks the first condition. For the second, one remarks that the path in $\mathrm{SO}(\mathbb{R}^2)$ which defines the vector $R_{g*}X$ is $(R_{g*}X)(t) = \{g^{-1} b(t) e_i\}_x$. It follows that

$$\begin{aligned}
\omega(R_{g*}X) &= -(g^{-1} b(t) b(0)^{-1} g)'(0)\\
&= -\left(\mathbf{Ad}_{g^{-1}}(b(t)b(0)^{-1})\right)'(0)\\
&= -Ad_{g^{-1}}(b(t)b(0)^{-1})'(0)\\
&= Ad_{g^{-1}}\omega(X).
\end{aligned} \tag{5.109}$$

□

Proposition 5.49.
The covariant derivative induced on M *by this connection is*

$$\nabla_X Y = X(Y). \tag{5.110}$$

Proof. In this demonstration, we will use the equivariant functions defined in 4.5.4. In order to compute $(\nabla_X Y)_x$, we have to use the definition of theorem 4.27. We first have to compute the horizontal lift of X. It is no difficult to see that $\overline{X}_{\{be_i\}_x}$ is given by the path

$$\overline{X}(t) = \{be_i\}_{X(t)}$$

if the vector field X is given by the path $X(t)$ in M. Indeed, it is trivial that $\omega(\overline{X}) = 0$, and

$$d\pi_* \overline{X} = \frac{d}{dt}\pi\{be_i\}_{X(t)}\Big|_{t=0} = \frac{d}{dt}X(t)\Big|_{t=0} = X.$$

Now, we compute $(\overline{X}\hat{s})(b)$ for $b = \{Se_i\}_x$. We begin using the basic definitions and notations:

$$(\overline{X}\hat{s})(b) = \overline{X}_b\hat{s} = \frac{d}{dt}\hat{s}(\overline{X}_b(t))\Big|_{t=0} = \frac{d}{dt}\hat{s}(\{Se_i\}_{X(t)})\Big|_{t=0}.$$

We can rewrite it with $\hat{Y}$ instead of $\hat{s}$. By construction (see (4.60)), if $b = \{Se_i\}_x$, $\hat{Y}(b) = S^{-1}(Y_x)$. Thus

$$(\overline{X}\hat{Y})(b) = \frac{d}{dt}S^{-1}(Y_{X(t)})\Big|_{t=0},$$

where, if $\{\overline{1}_i\}$ is a basis of $\mathbb{R}^m$, then S is

$$\begin{aligned} S\colon \mathbb{R}^m &\to T_{X(t)}M \\ v^i\overline{1}_i &\mapsto S^i_j v^j(\partial_j)_{X(t)} \end{aligned} \tag{5.111}$$

So if we write $Y_x = Y^i(x)\partial_i$, we have

$$S^{-1}(Y_{X(t)}) = (S^{-1})^i_j Y^j(X(t))\overline{1}_i$$

and

$$\frac{d}{dt}S^{-1}(Y_{X(t)})\Big|_{t=0} = (S^{-1})^i_j \frac{d}{dt}Y^j(X(t))\Big|_{t=0}\overline{1}_i = (S^{-1})^i_j X(Y^j)\overline{1}_i.$$

Since b is an isomorphism, we can apply b on both side of $\hat{X}(b) = b^{-1}(X_x)$, and take $\nabla_X Y$ instead of X:

$$(\nabla_X Y)(x) = b\big((S^{-1})^i_j X(Y^j)\overline{1}_i\big) = S^k_i (S^{-1})^i_j X(Y^j)(\partial_k)_x = X(Y^j)(\partial_j)_x = X(Y)_x. \tag{5.112}$$

□

From this and definition 4.70, we immediately conclude that our connection is torsion-free. In a certain manner, one can say that our covariant derivative is the usual one.

5.8.2 Construction of γ

Now, we construct the map γ of subsection 5.7. The first step is to define $\hat{a}_X\colon P \to \mathrm{End}\,(\Lambda W)$ by

$$\hat{a}_X(p) = \tilde{\rho}(\hat{X}_{\varphi(p)}).$$

Here, ΛW is the completely isotropic subspace of $(\mathbb{R}^2)^{\mathbb{C}}$ with euclidian metric; thus we can use the result of section 5.4.5. In particular, we know the representation $\tilde{\rho}$.

To see it more explicitly, we need the expression of $\hat{X}$. It is given in subsection 4.5.4: if b is the basis $\{be_i\}_x$, $\hat{Y}(b) = b^{-1}(Y_x)$. As $\varphi(b,x) = \{be_i\}_x$, we have

$$\hat{a}_X(b,x) = \tilde{\rho}(b^{-1}(X_x)).$$

The subsection 4.5.2.2 explains how to explicitly get $\gamma(X)$ with the definition $\gamma(X) = a_X$. If ψ is a section of $\mathcal{S}$ and $\psi(x) = [\xi, v]$, the general definition gives us $(a_X\psi)(x) = [\xi, \hat{a}_X(\xi)v]$ and in our particular case, if $\xi = (b,x)$, we get:

$$(\gamma(X)\psi)(x) = [\xi, \tilde{\rho}(b^{-1}(X_x))v]. \tag{5.113}$$

5.8.3 Covariant derivative on $\Gamma(\mathcal{S})$

Remember the spin structure of $\mathrm{SO}(\mathbb{R}^2)$: $\varphi(x,S) = \{Se_i\}_x$. We now construct the connection on $P = \mathbb{R}^2 \times \mathrm{SO}(2)$. It is defined by the 1-form $\tilde{\omega} = \varphi^*\omega$. If v is a vector of P, it is described by a path $v(t) = (x(t), b(t))$, then the path of $d\varphi(v)$ is $\{b(t)e_i\}_{x(t)}$ and $\tilde{\omega}(v) = \omega(d\varphi(v)) = -(b(t)b(0)^{-1})'(0)$.

The next step defining the Dirac operator is to find out an explicit form for the map $\tilde{\nabla}\colon \mathfrak{X}(M)\times\Gamma(\mathcal{S}) \to \Gamma(\mathcal{S})$. A section $s \in \Gamma(\mathcal{S})$ is a map $s\colon M \to \mathcal{S} = (\mathbb{R}^2 \times \mathrm{SO}(2)) \times_\rho \Lambda W$; it is defined by an equivariant function $\hat{s}\colon P \to \Lambda W$. In order to find the value of $(\tilde{\nabla}_X s)(x)$ for $X \in \mathfrak{X}(M)$, we use the definition

$$\widehat{\tilde{\nabla}_X s}(\xi) = \overline{X}_\xi(\hat{s})$$

where $\overline{X}$ is the horizontal lift in the sense of $\tilde{\omega}$. For the same reason as in the proof of proposition 5.110, $\overline{X}_{(b,x)}$ is given by the path $\overline{X}(t) = (b, X(t))$ where $X(t)$ is the path which defines X. So we have

$$\widehat{\tilde{\nabla}_X s}(\xi) = \overline{X}_{(b,x)}(\hat{s}) = \frac{d}{dt}\hat{s}(b, X(t))\Big|_{t=0}.$$

Remark that ΛW is a vector space; so for every $\alpha \in \Lambda W$, the identification $T_\alpha \Lambda W = \Lambda W$ is correct.

Our first form of $\tilde{\nabla}$ is

$$(\tilde{\nabla}_X s)(x) = \Big[\xi, \frac{d}{dt}\hat{s}(b, X(t))\Big|_{t=0}\Big],$$

but we can modify this in order to get simpler expressions. Remark that we have an equivalence class, so that we can always choose the element of the class such that $\xi = (\mathbb{1}, x)$. We define $\overline{s}\colon \mathbb{R}^2 \to \Lambda W$, $\overline{s}(v) = \hat{s}(\mathbb{1}, v)$. Our second and final form for $\tilde{\nabla}$ is:

$$(\tilde{\nabla}_X s)(x) = \Big[(\mathbb{1}, x), \frac{d}{dt}\overline{s}(X(t))\Big|_{t=0}\Big] \tag{5.114a}$$

$$= [(\mathbb{1}, x), X(\overline{s})], \tag{5.114b}$$

where $X(\overline{s})$ is well defined because $\overline{s}$ is a map from $\mathbb{R}^2$ into a vector space (namely: ΛW).

5.8.4 Dirac operator on the euclidian $\mathbb{R}^2$

We continue to write explicitly the definition (5.97). Putting together (5.113) and (5.114b), one finds

$$\gamma_x^\alpha(\tilde{\nabla}_{e_\beta} s)(x) = \gamma(e_{\alpha x})[\xi, e_\beta(\overline{s})] = [\xi, \tilde{\rho}(b^{-1}(e_{\alpha x}))e_\beta(\overline{s})]. \tag{5.115}$$

Here, $e_\beta = \partial_\beta$ and $b = \mathbb{1}$, then

$$\gamma_x^\alpha(\tilde{\nabla}_{e_\beta} s)(x) = [(\mathbb{1}, x), \tilde{\rho}(e_\alpha)\partial_\beta \overline{s}].$$

Now, the Dirac operator reads

$$(\mathcal{D}s)(x) = [(\mathbb{1}, x), \gamma^\alpha \partial_\alpha \overline{s}].$$

We can obtain a more compact expression by defining "Ys" and "As" when $s \in \Gamma(\mathcal{S})$, $Y \in \mathfrak{X}(\mathbb{R}^2)$ and $A \in \operatorname{End}\Lambda W$. The definitions are

$$(Ys)(x) = [(\mathbb{1}, x), (Y\overline{s})(x)],$$
$$(As)(x) = [(\mathbb{1}, x), A\overline{s}(x)].$$

With these conventions, one writes:

$$(\mathcal{D}s)(x) = \gamma^\alpha(\partial_\alpha s)(x).$$

This justifies the expression (5.3): $\mathcal{D} = \gamma^\alpha \partial_\alpha$ on flat spaces. With a good choice of basis of ΛW, the matrices γ^α are given by (5.62), and

$$\gamma^\alpha \partial_\alpha = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \partial_x - \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \partial_y.$$

If we identify $\mathbb{R}^2$ with $\mathbb{C}$ we have the following definitions:

$$\partial_z = \frac{1}{2}(\partial_x - i\partial_y), \qquad \partial_{\overline{z}} = \frac{1}{2}(\partial_x + i\partial_y),$$

so that

$$\mathcal{D} = \begin{pmatrix} 0 & -\partial_{\overline{z}} \\ \partial_z & 0 \end{pmatrix}.$$

5.9 Clifford algebras and Morita equivalence

Let $\mathfrak{A}$ be an algebra. An algebra $\mathcal{B}$ is said to be **Morita equivalent** to $\mathfrak{A}$ if $\mathcal{B} = \operatorname{End}_{\mathfrak{A}}(\mathcal{E})$ for some finite projective module $\mathcal{E}$ over $\mathfrak{A}$. The algebra $\mathfrak{A}$ is Morita equivalent to itself taking the trivial module $\mathcal{E} = \mathfrak{A}$.

We consider a manifold M of dimension $n = 2m$.

Problem and misunderstanding 25.
The two following statements are imprecise.

Proposition 5.50.
A module which implement a Morita equivalence between two C^*-algebras is finite projective.

Theorem 5.51 (Serre-Swan)**.**
If one of the two Morita equivalent is the continuous function space over a manifold $\mathfrak{A} = C(M)$, then the module which gives the Morita equivalence is the section of continuous sections of a vector bundle over M, $\mathcal{E} = \Gamma(E)$.

Furthermore, if $\mathfrak{A} = C(M)$ and $\mathcal{B} = \Gamma(\mathrm{Cl}(M))$, we have $\operatorname{End} E \simeq \mathrm{Cl}(M)$ as isomorphism of vector bundle. Since $\mathrm{Cl}\, M$ is of rank 2^n, $\operatorname{End} E$ has same rank and E_x has dimension $\sqrt{2^n} = 2^m$. So it is possible to choose the Clifford action in such a way that $\Gamma(E)$ is an irreducible Clifford module.

We often look at an anti-linear map $J\colon \Gamma(E) \to \Gamma(E)$ such that for all $\psi \in \Gamma(E)$

(i) $J(\psi f) = (J\psi)\overline{f}$ for all $f \in C(M)$,

(ii) $J(a\psi) = \epsilon(a)aJ\psi$ for all $a \in \Gamma^\infty(\mathrm{Cl}\, M)$.

How to define $a\psi$? We consider $\mathfrak{A} = C(M)$, $\mathcal{B} = \Gamma(\mathrm{Cl}\, M)$ and we define $\Gamma(E)$ is such a way that it implements a Morita equivalence between $\mathfrak{A}$ and $\mathcal{B}$; hence $\Gamma(E)$ is a $C(M)$-module. From dimensional considerations, we can define on $\Gamma(E)$ a Clifford module structure, i.e. a $C(M)$-linear

$$c\colon \Gamma(\mathrm{Cl}\, M) \to \operatorname{End}(\Gamma E), \tag{5.116}$$

hence $a\psi$ makes sense for any $a \in \Gamma^\infty(\mathrm{Cl}\, M)$ and $\psi \in \Gamma(E)$ with definition

$$(a\psi)(x) = \big(c(a)\psi\big)(x) = c(a(x))\psi(x) \tag{5.117}$$

Theorem 5.52.
Let (M, S, J) be a spin manifold of dimension n. There exists an unique connection

$$\nabla^S \colon \Gamma^\infty(S) \to \Gamma^\infty(S) \otimes \Omega^1(S)$$

such that

(i) $(\nabla^S\psi|\phi) + (\psi|\nabla^S\phi) = d(\psi|\phi)$,

(ii) $[\nabla^S, J] = 0$,

(iii) $\nabla^S\big(c(a)\psi\big) = c\big(\nabla a\big)\psi + c(a)\nabla^S\psi$ *for all* $a \in \mathrm{Cl}(M)$ *and* $\psi \in \Gamma^\infty(S)$.

In the latter, the action of $\Gamma^\infty(\mathrm{Cl}\, M)$ on $\Gamma^\infty(S)$ is induced from the action $c\colon \mathrm{Cl}(T_x^*M) \to \operatorname{End} S$. The ∇ which acts on a is the connection extended to $\Gamma^\infty(\mathrm{Cl}\, M)$ by virtue of Leibnitz rule $\nabla(uv) = \nabla(u)v + u\nabla(c)$.

Proof. No proof □

In this setting, we define

$$\begin{aligned} \hat{c}\colon \Gamma^\infty(S) \otimes \Gamma^\infty(\mathrm{Cl}\, M) &\to \Gamma^\infty(S) \\ \psi \otimes a &\mapsto c(a)\psi. \end{aligned} \tag{5.118}$$

Then we define the **Dirac operator** $\mathcal{D}\colon \Gamma^\infty(S) \to \Gamma^\infty(S)$,

$$\mathcal{D} = -i(\hat{c} \circ \nabla^S). \tag{5.119}$$

5.9.1 Example: quantum field theory

Let us show how does this operator gives back the usual Dirac operator of quantum field theory. Let M be a manifold and with two local basis $\{\partial_u\}$ and $\{\partial_\alpha\}$ of T_xM. The first one is the "natural" basis: $g(\partial_u, \partial_v) = g_{uv}$ has no particular properties while the second one is orthonormal $g(\partial_\alpha, \partial_\beta) = \delta_{\alpha\beta}$. The first dual basis is defined by $dx^\alpha\partial_\beta = \delta^\alpha_\beta$.

We write $\partial_\alpha = e^u_\alpha\partial_u$ and for the dual basis, $dx^\alpha = e^\alpha_u\, dx^u$. In order these definition to be coherent, we impose $dx^\alpha\partial_\beta = \delta^\alpha_\beta$:

$$dx^\alpha\partial_\beta = e^\alpha_u dx^u\big(e^v_\beta\partial_v\big) = e^\alpha_u e^v_\beta \delta^u_v = e^\alpha_u e^u_\beta. \tag{5.120}$$

We conclude that the **vielbein** (e^α_u) is the inverse of (e^v_β): $e^\alpha_u e^u_\beta = \delta^\alpha_\beta$. The vielbein are eventually complexes.

5.9.2 An other definition of the Dirac operator

Let us consider an orthonormal basis $\{e_a\}$ of M, i.e. on each $x \in M$,

$$g_x(e_a(x), e_b(x)) = \eta_{ab}.$$

This basis is related to a "natural" basis $\{\partial_\mu\}$ by

$$e_a = e_a^\mu \partial_\mu \tag{5.121}$$

where e_a^μ is called **vielbein** (here, they are more precisely n-beins). As far as metric is concerned we have

$$g^{\mu\nu} = e_a^\mu e_b^\nu \eta_{ab} \tag{5.122a}$$

$$\eta_{ab} = e_a^\mu e_b^\nu g_{\mu\nu}. \tag{5.122b}$$

If ∇ is the covariant derivative associated with g, we define the coefficients $\omega_{\mu a}^b$ by

$$\nabla_\mu e_a = \omega_{\mu a}^b e_b. \tag{5.123}$$

On the other hand, ∇ is related to the Christoffel symbols by

$$\nabla_\mu \partial_\nu = \Gamma_{\mu\nu}^\sigma \partial_\sigma. \tag{5.124}$$

Let $\mathrm{Cl}(M)$ be the Clifford module whose fibre is the Clifford complex algebra $\mathrm{Cl}(T_x^*M)^{\mathbb{C}}$. We consider $\Gamma(\mathrm{Cl}(M))$, the module of corresponding sections. It gives an algebra morphism

$$\begin{aligned}\gamma\colon \Gamma(\mathrm{Cl}(M)) &\to \mathfrak{B}(\mathscr{H})\\ dx^\mu &\mapsto \gamma^\mu(x) = \gamma^a e_a^\mu\end{aligned} \tag{5.125}$$

which can be extended to the whole Clifford algebra. One can choose matrices $\gamma^\mu(x)$ and γ^a to be hermitian; they satisfy

$$\gamma^\mu(x)\gamma^\nu(x) + \gamma^\nu(x)\gamma^\mu(x) = -2g(dx^\mu, dx^\nu) = -2g^{\mu\nu} \tag{5.126a}$$

$$\gamma^a\gamma^b + \gamma^b\gamma^a = -2\eta^{ab}. \tag{5.126b}$$

All this allow us to lift the Levi-Civita connection from the tangent bundle to the spinor bundle by defining

$$\nabla_\mu^S = \partial_\mu + \omega_\mu^S = \partial_\mu + \frac{1}{2}\omega_{\mu ab}\gamma^a\gamma^b. \tag{5.127}$$

The **Dirac operator** is then given by

$$\mathcal{D} = \gamma \circ \nabla$$

and can locally be written under the form

$$\mathcal{D} = \gamma^\mu(x)(\partial_\mu + \omega_\mu^S) = \gamma^a e_a^\mu(\partial_\mu + \omega_\mu^S). \tag{5.128}$$

5.10 Dirac operator on AdS_2

Why to compute Dirac operator on anti de Sitter spaces ? Let $M = AdS_2$ and $R = AN$ acts on M. Let $\mathcal{O}$ be an open orbit of $R \times M \to M$. In the specific case of AdS_2, we have $R = \mathcal{O} = R \cdot \vartheta$. In larger dimensions, there is a $\mathrm{SO}(1,n)$ which causes that the orbit is not exactly the acting group. It is

$$\mathcal{O} = \frac{R}{R \cap \mathrm{SO}(1,n)}.$$

5.10.1 Clifford algebra and spin group

As definition, we retain

$$\begin{aligned}AdS_2 &\equiv t^2 + u^2 - x^2 = 1\\ &= \frac{\mathrm{SO}(2,1)}{\mathrm{SO}(1,1)}.\end{aligned} \tag{5.129}$$

Let $V = \mathbb{R}^{1,1}$ and e_0, e_1 an orthonormal basis. We pose

$$f_0 = \frac{1}{2}(e_0 + e_1) \quad g_0 = \frac{1}{2}(e_0 - e_1)$$

and we define $\tilde{\rho}$ by

$$\tilde{\rho}(f_0)\alpha = f_0 \wedge \alpha \tag{5.130a}$$

$$\tilde{\rho}(g_0)\alpha = -i(g_0)\alpha \tag{5.130b}$$

where $\alpha \in \Lambda W$, W being the space spanned by f_0. More explicitly we have :

$$\tilde{\rho}(f_0)1 = f_0 \qquad \tilde{\rho}(f_0)f_0 = 0 \tag{5.131a}$$

$$\tilde{\rho}(g_0)1 = 0 \qquad \tilde{\rho}(g_0)f_0 = -\eta(f_0, g_0). \tag{5.131b}$$

As element of ΛW, f_0 stands for $\eta(f_0, .)$. If we choose the basis

$$1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad f_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

the matrices of $\tilde{\rho}$ are given by

$$\tilde{\rho}(e_0) = \begin{pmatrix} 0 & -1/2 \\ 1 & 0 \end{pmatrix}, \quad \tilde{\rho}(e_1) = \begin{pmatrix} 0 & 1/2 \\ 1 & 0 \end{pmatrix}.$$

Up to a change of basis,

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \gamma_0\gamma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and a general element of $Cl_{(1,1)}$ reads

$$x\gamma_0 + y\gamma_1 + u\mathbb{R} + v\gamma_0\gamma_1 = \begin{pmatrix} u+v & x-y \\ x+y & u-v \end{pmatrix}.$$

With the change of basis $e_1 \to ie_1$, we write it under a more simple form :

$$Cl_{(1,1)} \rightsquigarrow \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} \tag{5.132}$$

with $\alpha, \beta \in \mathbb{C}$. In particular, an element of V, i.e. a combination of γ_0 and γ_1 is

$$V \rightsquigarrow \begin{pmatrix} 0 & \xi \\ \bar{\xi} & 0 \end{pmatrix}. \tag{5.133}$$

Let

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad a = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad c = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$$

Let us now determine α, the extension of $-\operatorname{id}|_V$ into an automorphism and τ, the extension of $\operatorname{id}|_V$ into an anti-automorphism. We have $\alpha(b) = -b$, $\alpha(c) = -c$, $\tau(b) = b$ and $\tau(c) = c$. We find the others by virtue of relations $bc = -a$ and $b^2 = 1$. Finally

$$\begin{aligned} \alpha(1) &= 1 & \alpha(a) &= a \\ \alpha(b) &= -b & \alpha(c) &= -c \end{aligned} \tag{5.134}$$

and

$$\begin{aligned} \tau(1) &= 1 & \tau(a) &= -a \\ \tau(b) &= b & \tau(c) &= c. \end{aligned} \tag{5.135}$$

The condition for $s \in Cl_{(1,1)}$ to belongs to $\Gamma_{(1,1)}$ is that $\alpha(s)vs^{-1} \in V$ for all $v \in V$. If we consider $s = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}$, we have

$$\alpha(s) = \begin{pmatrix} \alpha & -\beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}, \text{ and } s^{-1} = \frac{1}{|\alpha|^2 - |\beta|^2} \begin{pmatrix} \bar{\alpha} & -\beta \\ -\bar{\beta} & \alpha \end{pmatrix}.$$

If we impose $\alpha(s)vs^{-1}$ to be of the form $\begin{pmatrix} 0 & \eta \\ \bar{\eta} & 0 \end{pmatrix}$ for all v of the form $\begin{pmatrix} 0 & \xi \\ \bar{\xi} & 0 \end{pmatrix}$, we find $\operatorname{Re}(\bar{\alpha}\beta\bar{\xi}) = 0$ and the $\bar{\alpha}\beta = 0$. So generators of $\Gamma_{(1,1)}$ are

$$\Gamma_{(1,1)} \rightsquigarrow \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix}, \quad \begin{pmatrix} 0 & \beta \\ \bar{\beta} & 0 \end{pmatrix}. \tag{5.136}$$

Elements of $Spin_{(1,1)}$ are elements of $\Gamma^+_{(1,1)}$ such that $\tau(s) = s^{-1}$. So

$$Spin_{(1,1)} \rightsquigarrow \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix}, \text{ such that } |\alpha|^2 = 1. \tag{5.137}$$

We recognize $Spin_{(1,1)} = U(1)$.

$\boxed{\textbf{Problem and misunderstanding}}$ **26.**
This is wrong: in fact $\mathrm{Spin}(1,1) \neq U(1)$.

5.10.2 Relation between $SU(1,1)$ and $\mathrm{SO}(2,1)$

A general matrix of $SU(1,1)$ is

$$U = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}, \quad U^{-1} = \begin{pmatrix} \bar{\alpha} & -\beta \\ -\bar{\beta} & \alpha \end{pmatrix}$$

with $|\alpha|^2 - |\beta|^2 = 1$. They are matrices which fulfil $\det U = 1$ and $U^+ g = gU^{-1}$. If we denote by V the space of matrices of the form $(r,z) = \begin{pmatrix} r & \bar{z} \\ z & r \end{pmatrix}$ with $r \in \mathbb{R}$ and $z \in \mathbb{C}$, we have a bijection $\psi \colon \mathbb{R}^{1,1} \to V$ given by

$$\begin{pmatrix} u \\ t \\ x \end{pmatrix} \mapsto \begin{pmatrix} x & t - iu \\ t + iu & x \end{pmatrix}.$$

It becomes an isometry if we pose $\|(r,z)\| = z\bar{z} - r^2 = -\det(r,z)$. The group $SU(1,1)$ has an isometric action on V given by

$$Uv = UvU^\dagger.$$

We immediately remark that $Uv = (-U)v$. We define

$$\begin{aligned} &T \colon SU(1,1) \to \mathrm{SO}(2,1) \\ &T(U)\begin{pmatrix} u \\ t \\ x \end{pmatrix} = \psi^{-1}\big(U\psi\begin{pmatrix} u \\ t \\ x \end{pmatrix}U^\dagger\big). \end{aligned} \tag{5.138}$$

Now we want to know when $T(U) = T(\tilde{U})$. Using the fact that $U^{-1} = gU^\dagger g$ in the condition $UvU^\dagger = \tilde{U}v\tilde{U}^\dagger$, we find

$$VvV^\dagger = v$$

with $V = \tilde{U}^{-1}U$. Then imposing

$$\begin{pmatrix} r & \bar{z} \\ z & r \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}\begin{pmatrix} r & \bar{z} \\ z & r \end{pmatrix}\begin{pmatrix} \bar{\alpha} & \beta \\ \bar{\beta} & \alpha \end{pmatrix},$$

we find $T(U) = T(\tilde{U}) \Leftrightarrow \tilde{U} = \pm U$. We have

$$T\begin{pmatrix} i & \\ & -i \end{pmatrix} = \begin{pmatrix} -1 & & \\ & -1 & \\ & & 1 \end{pmatrix}.$$

The map $T \colon SU(1,1) \to \mathrm{SO}(2,1)$ is a double covering.

We are now going to explicitly compute the map T. First :

$$\begin{aligned} &\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix}\begin{pmatrix} r & \bar{z} \\ z & r \end{pmatrix}\begin{pmatrix} \bar{\alpha} & \beta \\ \bar{\beta} & \alpha \end{pmatrix} = \\ &\qquad\begin{pmatrix} \bar{\alpha}(\alpha r + \beta z) + \bar{\beta}(\alpha\bar{z} + \beta r) & \beta(\alpha r + \beta z) + \alpha(\alpha\bar{z} + \beta r) \\ \bar{\alpha}(\bar{\beta} + \bar{\alpha}z) + \bar{\beta}(\bar{\beta}\bar{z} + \bar{\alpha}r) & \beta(\bar{\beta}r + \bar{\alpha}z) + \alpha(\bar{\beta}\bar{z} + \bar{\alpha}r) \end{pmatrix}. \end{aligned}$$

When we pose $z = 0$ and $r = 1$, i.e., when we look at $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, we find

$$\begin{pmatrix} \bar{\alpha}\alpha + \bar{\beta}\beta & 2\beta\alpha \\ 2\bar{\alpha}\bar{\beta} & \beta\bar{\beta} + \alpha\bar{\alpha} \end{pmatrix}$$

which corresponds to $x = \bar{\alpha}\alpha$, $t - iu = 2\alpha\beta$ and $t + iu = 2\bar{\alpha}\bar{\beta}$. We conclude that

$$T\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} = \begin{pmatrix} . & . & i(\alpha\beta - \bar{\alpha}\bar{\beta}) \\ . & . & \alpha\beta + \bar{\alpha}\bar{\beta} \\ . & . & \alpha\bar{\alpha} + \beta\bar{\beta} \end{pmatrix}.$$

Similar computations lead to

$$T\begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} = \begin{pmatrix} \frac{\bar{\alpha}^2+\alpha^2-\beta^2-\bar{\beta}^2}{2} & \frac{i}{2}(\alpha^2 - \bar{\alpha}^2 + \beta^2 - \bar{\beta}^2) & i(\alpha\beta - \bar{\alpha}\bar{\beta}) \\ \frac{i}{2}(\beta^2 - \bar{\beta}^2 + \bar{\alpha}^2 - \alpha^2) & \frac{1}{2}(\alpha^2 + \bar{\alpha}^2 + \beta^2 + \bar{\beta}^2) & \alpha\beta + \bar{\alpha}\bar{\beta} \\ i(\bar{\alpha}\beta - \bar{\beta}\alpha) & \bar{\alpha}\beta + \bar{\beta}\alpha & \alpha\bar{\alpha} + \beta\bar{\beta} \end{pmatrix} \tag{5.139}$$

5.10.3 Spin structure on AdS_2

We are going to build elements of the following spin structure:

$$\begin{tikzcd} \mathrm{Spin}(1,1) \arrow[r, squiggly] & SU(1,1) \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & \mathrm{SO}(AdS_2) \arrow[dl, "p"] & \mathrm{SO}(2,1) \arrow[l, squiggly] \\ & & AdS_2 & & \end{tikzcd}$$

First let $\{e_t, e_u, e_x\}$ be a basis of $\mathbb{R}^{1,1}$ with $e_t \in AdS_2$, $e_t\cdot e_t = e_u\cdot e_u = -e_x\cdot e_x = -1$. We suppose that e_u and e_x span tangent space at e_t. Let T be a representation of $\mathrm{SO}(1,1)$ on $\mathbb{R}^{2,1}$ which leaves e_t unchanged: $T(A)e_t = e_t$ for all $A \in \mathrm{SO}(1,1)$. To each element $B \in \mathrm{SO}(AdS_2)$, one can associate an element of $B' \in \mathrm{SO}(AdS_2)$ such that B has the form

$$B = \{B'e_u, B'e_x\}_{B'e_t}. \tag{5.140}$$

We define

$$p(B) = B'e_t.$$

Now the action of $A \in \mathrm{SO}(1,1)$ on $B \in \mathrm{SO}(AdS_2)$ is defined, if B has the form (5.140), by

$$B \cdot A = \{T(A)B'e_u, T(A)B'e_x\}_{B'e_t}. \tag{5.141}$$

The map $\varphi\colon SU(1,1) \to \mathrm{SO}(AdS_2)$ is given by

$$\big(\varphi(U)\big)' = (T \circ S)(U), \tag{5.142}$$

and the projection $\pi\colon SU(1,1) \to AdS_2$, $\pi = p \circ \varphi$.

The group $\mathrm{Spin}(1,1)$ must act on $SU(1,1)$; we define

$$U \cdot s = \varphi^{-1}\big(\varphi(U) \cdot \chi(s)\big). \tag{5.143}$$

We have $\pi(U \cdot s) = \pi(U)$ because

$$\pi(U \cdot s) = p\big(\varphi(U)\cdot\chi(s)\big), = \big[\varphi(U)\cdot\chi(s)\big]'e_t = \varphi(U) \circ T(\chi(s))e_t = \varphi(U)'e_t = \pi(U).$$

We have used the fact that $\chi(s) \in \mathrm{SO}(1,1)$ and that, therefore, $(T \circ \chi)(s)e_t = e_t$.

5.10.4 Spinor bundle and connection

We define $S = \Lambda W$ where W is the (one dimensional) space spanned by f_0 and we define

$$\mathcal{S} = SU(1,1) \times_\rho S \tag{5.144}$$

where $\rho\colon Spin_{(1,1)} \times \Lambda W \to \Lambda W$ is the representation of $Spin_{(1,1)}$ on $SU(1,1)$ given by

$$\rho(s, \alpha) = \tilde{\rho}(s)\alpha. \tag{5.145}$$

Recall that α is either a scalar either a multiple of f_0. The equivalence relation which arises in equation (5.144) is

$$(U, \alpha) \sim (U \cdot s, \rho(s^{-1})\alpha). \tag{5.146}$$

The projection is

$$\pi_{\mathcal{S}}[(U, \alpha)] = \pi(U).$$

For the connection on $\mathrm{SO}(AdS_2)$, we want that horizontal vector are tangent vectors to curves formed by parallel transport. In other word, a path

$$B(s) = \{B'(s)e_u, B'(s)e_x\}_{B'(s)e_t}$$

has horizontal tangent vector if $B'(s)e_i$ $(i = u, x)$ is a parallel transport of $B'(0)e_i$ along the curve $B'(s)e_e$ on AdS_2. Here, $B'(s)$ denotes the matrix of $\mathrm{SO}(2,1)$ associated with the basis $B(s)$: the prime doesn't denotes a derivation. Let us define the $so(1,1)$ valued connection 1-form which corresponds to this intuition. We consider $b_i(s)$ the parallel transported along the curve $B'(s)e_t$ of $B'(0)e_i$, and $A(s)$, the matrix of $\mathrm{SO}(1,1)$ such that $A(s)B'(s)e_i = b_i(s)$ $(i = u, x)$. The definition is

$$\omega(\dot{B}) = \frac{d}{ds}\Big[A(s)\Big]_{s=0}.$$

Proposition 5.53.
It is a connection 1*-form.*

Proof. First we consider a fundamental vector field

$$X_B^* = \frac{d}{dt}\Big[B \cdot e^{-tX}\Big]_{t=0} = \frac{d}{dt}\Big[\{T(e^{-tX})B'e_u, T(e^{-tX})B'e_x\}_{B'e_t}\Big]_{t=0}.$$

The path in AdS_2 on which this path in $\mathrm{SO}(AdS_2)$ is build is constant: it is $B'e_t$. So the parallel transport is constant and the path $A(s)$ is given by

$$A(s)T(e^{-tX})B'e_u = B'e_u$$

and $\omega(X_B^*) = X$.

It remains to be proved that for all $B \in \mathrm{SO}(AdS_2)$, $g \in \mathrm{SO}(1,1)$ and $X \in T_B\,\mathrm{SO}(AdS_2)$,

$$\omega\big((dR_g)_B X\big) = \mathrm{Ad}(g^{-1})\omega_B(X). \tag{5.147}$$

We give X by the path

$$X(s) = \{B'(s)e_u, B'(s)e_x\}_{B'(s)e_t}.$$

The differential dR_g gives rise to the new path

$$(dR_g X)(s) = \{gB'(s)e_u, gB'(s)e_x\}_{B'(s)e_t}.$$

Let $b_i(s)$ be the parallel transport of $B'(0)e_i$ $(i = u, x)$ along the path $B'(s)e_t$. We have to compute $\omega_B(X)$ with $A(s)$ defined by $A(s)B'(s)e_i = b_i$. The parallel transport of $gB'(0)e_i$ is given by gb_i. Therefore $\omega(dR_gX)$ is given by the path $A^g(s)$ which satisfies $A^g(s)gB'(s)e_i = gA(s)B'(s)e_i$. So

$$A^g(s) = gA(s)g^{-1}$$

and

$$\frac{d}{ds}\Big[A^g(s)\Big]_{s=0} = \mathrm{Ad}(g)\omega(X).$$

□

5.10.5 Clifford algebra $(1,1)$

We consider the left invariant vector fields

$$\tilde{e}_J(r_0) = \frac{d}{ds}\Big[r_0 e^{-sJ}\Big]_{s=0} = -r_0 J \tag{5.148a}$$

$$\tilde{e}_L(r_0) = \frac{d}{ds}\Big[r_0 e^{-sL}\Big]_{s=0} = -r_0 L. \tag{5.148b}$$

More precisely, we consider the vectors given by action of these matrices on the "base point" $\begin{pmatrix}0\\1\\0\end{pmatrix}$. Hence

$$\tilde{e}_J(r_0) = -r_0\begin{pmatrix}0\\0\\1\end{pmatrix}, \quad \tilde{e}_L(r_0) = -r_0\begin{pmatrix}1\\0\\0\end{pmatrix} \tag{5.149}$$

and

$$g = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$$

Remark that this metric is constant (it does not depend on r_0) because r_0 is an isometry. For this reason, we now turn our attention to Clifford algebra and spin group for $V = \mathbb{R}^{1,1}$. Following matrices fulfill relation (5.30)

$$\gamma_J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma_L = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

The complete Clifford algebra has the following matrices too :

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \gamma_{JL} = \begin{pmatrix} -1 & \\ & 1 \end{pmatrix}.$$

The Clifford algebra is nothing else than $GL(2, \mathbb{R})$, the set of all real 2×2 matrices. From definitions, one can check that

$$\begin{aligned} \alpha(J) &= -J & \tau(J) &= J \\ \alpha(L) &= -L & \tau(L) &= L \\ \alpha(JL) &= JL & \tau(JL) &= -JL \\ \alpha(1) &= 1 & \tau(1) &= 1 \end{aligned}$$

Inverse and α of general element in $\mathrm{Cl}(1,1)$ are given by

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix}^{-1} = \frac{1}{ps - qr} \begin{pmatrix} s & -q \\ -r & p \end{pmatrix}, \quad \alpha \begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} p & -q \\ -r & s \end{pmatrix}.$$

A general element in $\mathbb{R}^{1,1}$ is $\begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}$ with $\alpha, \beta \in \mathbb{R}$, so the condition to belongs to $\Gamma(1,1)$ is that

$$\frac{1}{ps - qr} \begin{pmatrix} p & -q \\ -r & s \end{pmatrix} \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix} \begin{pmatrix} s & -q \\ -r & p \end{pmatrix}$$

belongs to $\mathbb{R}^{1,1}$ for all α and β. It requires, among others, that $qs\beta - rp\alpha = 0$ for all α and β. Hence $qs = rp = 0$, but the alternatives $p = r = 0$ and $q = s = 0$ are ruled out because we want the determinant $ps - qr$ to be non zero. Therefore, $\Gamma(1,1)$ is generated by

$$\Gamma(1,1) \rightsquigarrow \begin{pmatrix} p & 0 \\ 0 & s \end{pmatrix}, \begin{pmatrix} 0 & q \\ r & 0 \end{pmatrix}.$$

The latter belongs to $\mathbb{R}^{1,1}$, so

$$\Gamma^+(1,1) \rightsquigarrow \begin{pmatrix} z + c & 0 \\ 0 & z - c \end{pmatrix} = z\mathbb{1} + c\gamma_{JL}.$$

From

$$\tau(z\mathbb{1} + c\gamma_{JL}) = z\mathbb{1} - c\gamma_{JL},$$

elements in $\mathrm{Spin}(1,1)$ are subject to the relation

$$\begin{pmatrix} s & 0 \\ 0 & p \end{pmatrix} = \frac{1}{ps} \begin{pmatrix} s & 0 \\ 0 & p \end{pmatrix}.$$

As consequence, we find

$$\mathrm{Spin}(1,1) = \mathbb{R}_0 \rightsquigarrow \begin{pmatrix} 1/p & \\ & p \end{pmatrix}. \tag{5.150}$$

If we put (see decomposition (**??**))

$$A = \begin{pmatrix} e^a & \\ & e^{-a} \end{pmatrix},$$

we have $\mathrm{Spin}(1,1) = A \times \mathbb{Z}_2$. Let us check that $\mathrm{Spin}(1,1)$ is a double covering of $\mathrm{SO}_0(1,1)$. We know that

$$\mathrm{SO}(1,1) = \begin{pmatrix} \cosh\xi & \sinh\xi \\ \sinh\xi & \cosh\xi \end{pmatrix} \times \mathbb{Z}_2$$

while $\mathrm{SO}_0(1,1)$ is

$$\mathrm{SO}_0(1,1) = \begin{pmatrix} \cosh\xi & \sinh\xi \\ \sinh\xi & \cosh\xi \end{pmatrix} = \mathbb{R}.$$

This structure of SO$(1,1)$ comes from the fact (true for all SO$(1,n)$) that $|\Lambda^0_0| \geqslant 1$ when Λ is a Lorentz transformation. So $\mathbb{1}$ and $-\mathbb{1}$ cannot belong to the same connected component. Note that $\cosh\xi \geqslant 1$. We see intuitively how to cover two times $\mathbb{R}$ with $\mathbb{R}_0$. Let us see how the map χ does that. From definition, $\chi(x)y = \alpha(x)yx^{-1}$, so it is easy to see that

$$\chi(1) = \chi(-1) = \mathrm{id}\,|_{\mathbb{R}^{1,1}}$$

5.10.6 Parallel transport

We have a connection on the frame bundle of AdS_2 and we wan to lift the vectors $\tilde{e}_J$ and $\tilde{e}_L$, i.e. we consider a point

$$\xi_0 = (r_0, v_1, v_2) \in \mathrm{SO}(AdS_2)$$

where v_1 and v_2 form an orthonormal (in the sense of g) basis of $T_{r_0}AdS_2$. Then we have to find a path $s \to \xi(s)$ in SO(AdS_2) such that $\xi(0) = \xi_0$, $\omega(\xi'(0)) = 0$ and $dp\xi'(0) = \tilde{e}_a$. The latter condition allows us to compute $r(s)$ in the expression

$$\xi(s) = (r(s), v_1(s), v_2(s)),$$

namely, $r(s)$ is the path of $\tilde{e}_a$. The condition to be horizontal imposes that vectors $v_i(s)$ are parallel transport of v_i along $\tilde{e}_a$. So we have to compute the different $T_a(\tilde{e}_b)(s)$ which is the parallel transported of $\tilde{e}_b$ along the path of $\tilde{e}_a$ at a distance s; this is an element of $T_{\tilde{e}_a(s)}AdS_2$. It will be decomposed in the basis

$$\begin{aligned} \tilde{e}_J\big(\tilde{e}_a(s)\big) &= -r_0 e^{-sa} J \\ \tilde{e}_L\big(\tilde{e}_a(s)\big) &= -r_0 e^{-sJ} L. \end{aligned}$$

where we imply the action on the base point $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. For notational simplicity, from now we write $a(s)$ instead of $\tilde{e}_a(s)$. Various products are easy to compute; for example

$$\tilde{e}_J(J(s)) \cdot \tilde{e}_L(J(s)) = r_0 e^{-sJ} J \cdot r_0 e^{-sJ} L = J \cdot L = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0$$

because $r_0 e^{-sJ}$ is an isometry. In general :

$$\tilde{e}_a\big(c(s)\big) \cdot \tilde{e}_b\big(c(s)\big) = a \cdot b$$

Now we pose in general

$$T_a(\tilde{e}_b)(s) = \alpha(s)\tilde{e}_b\big(a(s)\big) + \beta(s)\tilde{e}_L\big(a(s)\big),$$

and we want to find the (real valued) functions α and β. Parallel transport fulfils two conditions: the norm and the angle with the path are constant. This leads us to two conditions :

$$T_a\big(\tilde{e}_b(s)\big) \cdot T_a\big(\tilde{e}_b(s)\big) = b \cdot b \tag{5.151a}$$
$$T_a\big(\tilde{e}_b(s)\big) \cdot \tilde{e}_a\big(a(s)\big) = b \cdot a. \tag{5.151b}$$

These equations extends to

$$\beta(s)^2 - \alpha(s)^2 = b \cdot b \tag{5.152a}$$
$$\alpha(s)J \cdot a + \beta(s)L \cdot a = b \cdot a. \tag{5.152b}$$

There are four cases to be considered following that $a = J, L$ and $b = J, L$. The result is that

$$T_a\big(\tilde{e}_b\big) = \tilde{e}_b, \tag{5.153}$$

in other terms, the vectors $\tilde{e}_J$ and $\tilde{e}_L$ are not only parallel vector fields, but each is parallel along the path of the other.

5.10.7 Covariant derivative

We will give the horizontal lift of $\tilde e_a$ at point

$$\xi(0)=\{B_1^b\tilde e_b, B_2^c\tilde e_c\}_{r_0e_t}$$

under the form of the path

$$\xi(s)=\{B_1^b\tilde e_b\big(a(s)\big), B_2^c\tilde e_c\big(a(s)\big)\}_{\tilde e_a(s)}.$$

We create a connection on the spinor bundle from the connexion via the formula

$$\widehat{\nabla^E_X\psi}(\xi)=\overline{X}_\xi(\hat\psi).$$

In our case, we take $\psi\colon M\to\mathcal{S}$, or $\hat\psi\colon SU(1,1)\to\Lambda W$ such that

$$\hat\psi(U\cdot g)=\rho(g^{-1})\hat\psi(U).$$

Since $\tilde\omega=\varphi^*\omega$, we have $\tilde\omega(X)=\omega(d\varphi X)$ and

$$\overline{e}_{a\xi_0}=\varphi^{-1}\big(\tilde e_a(s),\ldots\big).$$

Therefore

$$\widehat{\nabla_a\psi}(\xi_0)=\frac{d}{ds}\Big[(\hat\psi\circ\varphi^{-1})\{B_i^c\tilde e_c\big(a(s)\big)\}_{\tilde e_a(s)}\Big]_{s=0} \tag{5.154}$$

where φ is defined by

$$\varphi(U)=\{U\tilde e_J, U\tilde e_L\}_{Ur_0e_t}.$$

We have to find

$$\varphi^{-1}\{B_i^c\tilde e_c\big(a(s)\big)\}_{\tilde e_a}. \tag{5.155}$$

Before to write down the inverse of φ, let us perform some computations.

$$J=\begin{pmatrix} & 0 & \\ 0 & 0 & 1\\ & 1 & \end{pmatrix},\quad L=\begin{pmatrix}0 & 1 & 1\\ -1 & & \\ 1 & & \end{pmatrix},$$

and as far as we only wants to compute derivatives, we can write the exponentials as

$$e^{sJ}=\mathbb{1}+sJ=\begin{pmatrix}1 & & \\ & 1 & s\\ & s & 1\end{pmatrix} \tag{5.156}$$

$$e^{sL}=\mathbb{1}+sL=\begin{pmatrix}1 & s & s\\ -s & 1 & 0\\ s & 0 & 1\end{pmatrix}. \tag{5.157}$$

The path are given by

$$\tilde e_a(s)=r_0e^{-sa}\begin{pmatrix}0\\1\\0\end{pmatrix}, \tag{5.158}$$

in particular

$$\tilde e_J(s)=r_0\begin{pmatrix}0\\1\\-s\end{pmatrix},\qquad \tilde e_L(s)=r_0\begin{pmatrix}-s\\1\\0\end{pmatrix}. \tag{5.159}$$

For the various $\tilde e_b\big(a(s)\big)$, we have

$$\tilde e_b\big(a(s)\big)=\frac{d}{dt}\Big[a(s)e^{-tb}\Big]_{t=0}\begin{pmatrix}0\\1\\0\end{pmatrix}=\frac{d}{dt}\Big[r_0e^{-sa}e^{-tb}\Big]_{t=0}\begin{pmatrix}0\\1\\0\end{pmatrix}=-r_0e^{-sa}b\begin{pmatrix}0\\1\\0\end{pmatrix}. \tag{5.160}$$

Results are

$$\tilde e_J\big(J(s)\big)=-r_0\begin{pmatrix}0\\-s\\1\end{pmatrix}\qquad \tilde e_J\big(L(s)\big)=-r_0\begin{pmatrix}-s\\0\\1\end{pmatrix} \tag{5.161a}$$

$$\tilde e_L\big(J(s)\big)=-r_0\begin{pmatrix}1\\0\\0\end{pmatrix}\qquad \tilde e_L\big(L(s)\big)=-r_0\begin{pmatrix}1\\s\\-s\end{pmatrix}. \tag{5.161b}$$

We finally have to know that

$$B^c\tilde{e}_c\big(J(s)\big)=-r_0\begin{pmatrix}B^L\\-sB^J\\B^J\end{pmatrix},\quad B^c\tilde{e}_c\big(L(s)\big)=-r_0\begin{pmatrix}-sB^J+B^L\\sB^L\\B^J-sB^L\end{pmatrix}.$$

Following equation (5.155), in order to write down $\widehat{\nabla_a\psi}$, we have to find $U(s)\in SU(1,1)$ such that

(i) $Ur_0e_t=\tilde{e}_a(s)$,

(ii) $U\tilde{e}_J=B_1^c\tilde{e}_c\big(a(s)\big)$,

(iii) $U\tilde{e}_L=B_2^c\tilde{e}_c\big(a(s)\big)$.

If $\overline{f}$ and $\overline{g}$ are vectors, solutions in U of equation $Ur_0\overline{f}=r_0\overline{g}$ are $U=\mathbf{Ad}(r_0)B$ where B fulfils $B\overline{f}=\overline{g}$. In the case of $a=J$, the three conditions successively give

$$U=\mathbf{Ad}(r_0)\begin{pmatrix}.&0&.\\.&1&.\\.&-s&.\end{pmatrix}\tag{5.162a}$$

$$U=\mathbf{Ad}(r_0)\begin{pmatrix}.&.&B_1^L\\.&.&-sB_1^J\\.&.&B_1^J\end{pmatrix}\tag{5.162b}$$

$$U=\mathbf{Ad}(r_0)\begin{pmatrix}-B_2^L&.&.\\sB_2^J&.&.\\B_2^J&.&.\end{pmatrix}.\tag{5.162c}$$

Putting all together in equation (5.154) we find

$$\begin{aligned}\widehat{\nabla_J\psi}(\xi_0)&=\frac{d}{ds}\hat{\psi}\,\mathbf{Ad}(r_0)\begin{pmatrix}-B_2^L&0&B_1^L\\sB_2^J&1&-sB_1^J\\B_2^J&-s&B_1^J\end{pmatrix}\\&=d\hat{\psi}\,\mathbf{Ad}(r_0)\begin{pmatrix}0&0&0\\B_2^J&0&-B_1^J\\0&-1&0\end{pmatrix}.\end{aligned}\tag{5.163}$$

The same with L instead of J leads to

$$\widehat{\nabla_L\psi}(\xi_0)=d\hat{\psi}\,\mathbf{Ad}(r_0)\begin{pmatrix}-B_2^J&-1&B_1^J\\B_2^L&0&B_1^L\\-B_2^L&0&-B_1^L\end{pmatrix}.\tag{5.164}$$

However it should be shocking to get 3×3 matrices in $SU(1,1)$: we had abused between $\mathrm{SO}(2,1)$ and $SU(1,1)$.

5.10.8 Another way to write a section (wrong way to do)

The equivariant function $\hat{\psi}\colon SU(1,1)\to\Lambda W$ fulfills

$$\hat{\psi}(U\cdot g)=\rho(g^{-1})\hat{\psi}(U)$$

for all $g\in\mathrm{Spin}(1,1)$; in particular with $g=-\mathbb{1}$,

$$\hat{\psi}(-U)=-\hat{\psi}(U).\tag{5.165}$$

This gives the idea that it is not impossible to define $\hat{\psi}$ from its projection on $\mathrm{SO}(2,1)$: we want to get $\tilde{\psi}\colon\mathrm{SO}(2,1)\to\Lambda W$ and define

$$\hat{\psi}(U)=\tilde{\psi}\big(T(U)\big).$$

More precisely, we parametrize $SU(1,1)$ by α and β such that $|\alpha|^2-|\beta|^2=1$. Then we divide $SU(1,1)$ into two parts: $\alpha=x+iy$ is green when $x>0$ and when $x=0$, $y<0$; α is red when $x<0$ and when $x=0$, $y>0$. When $\alpha=0$, we classify following β in the same way. The result is that U is green if and only if $-U$ is red. For a map $\tilde{\psi}\colon\mathrm{SO}(2,1)\to\Lambda W$, we define

$$\hat{\psi}(U)=\begin{cases}\tilde{\psi}\big(T(U)\big)&\text{if }U\text{ is green}\\-\tilde{\psi}\big(T(U)\big)&\text{if }U\text{ is red}\end{cases}\tag{5.166}$$

We define $T^{-1}\colon \mathrm{SO}(2,1) \to SU(1,1)$ as follows: $T^{-1}(A)$ is the green element of $SU(1,1)$ whose image by T is A. In any cases we have

$$\hat{\psi} \circ T^{-1} = \tilde{\psi}.$$

The meaning of equations (5.163) and (5.164) is that $\mathbf{Ad}(r_0)$ is a matrix whose inverse image by T should be given to $\hat{\psi}$; the difficulty is to know which of the two. When U_0 is green,

$$\begin{aligned}\widehat{\nabla_a\psi}(U_0) &= \frac{d}{ds}\Big[(\hat{\psi}\circ T^{-1})\,\mathbf{Ad}(r_0)\Big(\cdots\Big)\Big]_{s=0}\\ &= \frac{d}{ds}\Big[\tilde{\psi}\,\mathbf{Ad}(r_0)\Big(\cdots\Big)\Big]_{s=0},\end{aligned}$$

while when U_0 is red,

$$\widehat{\nabla_a\psi}(U_0) = -\frac{d}{ds}\Big[\tilde{\psi}\,\mathbf{Ad}(r_0)\Big(\cdots\Big)\Big]_{s=0}.$$

These two show that

$$\widetilde{\nabla_a\psi}\big(T(U_0)\big) = \frac{d}{ds}\Big[\tilde{\psi}\,\mathbf{Ad}(r_0)\Big(\cdots\Big)\Big]_{s=0} \tag{5.167}$$

All this is only proved in the interior of the green and red regions so that the path $U(s)$ keeps on only one region.

5.10.9 Once again

We see AdS_2 as[7] $\mathcal{O} = \mathrm{Ad}(G)H$ and we consider a base point $o = \mathrm{Ad}(k_0)H$ with $G = SL(2,\mathbb{R}) = ANK$. Let the principal bundle

$$\begin{tikzcd} A \arrow[r, squiggly, "R"] & G \arrow[d, "\pi"] \\ & \mathcal{O} \end{tikzcd}$$

with A acting on G by $(a,g) \mapsto ga$ and the projection

$$\pi(rk_0a) = \mathrm{Ad}(rk_0a)H. \tag{5.168}$$

where $r \in R$ and $a \in A$. More precisely, the principal bundle we look at is

$$\begin{tikzcd} A \arrow[r, squiggly, "R"] & \mathcal{U}_G \arrow[d, "\pi"] \\ & \mathcal{U} \end{tikzcd} \tag{5.169}$$

where $\mathcal{U}_G = Rk_0A$ and $\mathcal{U} = \pi(\mathcal{U}_G) = \mathrm{Ad}(Rk_0A)H = \mathrm{Ad}(Rk_0)H = \mathrm{Ad}(R)o$. The $\mathcal{U}_G$ is so defined in order to be the π^{-1} of an orbit $\mathcal{U} = \mathrm{Ad}(R)o$.

We have a manifold isomorphism $R \simeq \mathcal{U}$ given by

$$\phi\colon r \to \mathrm{Ad}(r)o.$$

How to see a left invariant vector field on R *via* this identification ?

$$d\phi\tilde{X}_r = d\phi\frac{d}{dt}\Big[re^{tX}\Big]_{t=0} = \frac{d}{dt}\Big[\mathrm{Ad}(r)\,\mathrm{Ad}(e^{tX})o\Big]_{t=0}.$$

This leads us to consider the following field for $X \in \mathcal{R}$. We define $\xi_X(rk_0a) \in T_{rk_0a}\mathcal{U}_G$,

$$\xi_X(rk_0a) = \frac{d}{dt}\Big[re^{tX}k_0a\Big]_{t=0}. \tag{5.170}$$

Let's see the projection :

$$\begin{aligned}d\pi_{rk_0a}\xi_X(rk_0a) &= \frac{d}{dt}\Big[\pi(re^{tX}k_0a)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\mathrm{Ad}(re^{tX}k_0a)H\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\mathrm{Ad}(re^{tX})o\Big]_{t=0}.\end{aligned}$$

[7]Here, $G = SL(2,\mathbb{R})$

This gives us the idea to define $X^\sharp \in T_{\mathrm{Ad}(rk_0a)H}\mathcal{U} = T_{\pi(rk_0a)}\mathcal{U}$ by

$$X^\sharp_{rk_0a} = \frac{d}{dt}\Big[\operatorname{Ad}(re^{tX})o\Big]_{t=0}, \tag{5.171}$$

which is a good definition because $\pi(rk_0a) = \pi(r'k_0a')$ only when $r = r'$. We put the following connection on $\mathcal{U}_G$:

$$\alpha_{rk_0a}(\Sigma) = \Big[\big(dL^{-1}_{rk_0a}\big)_{rk_0a}\Sigma\Big]_{\mathcal{A}}. \tag{5.172}$$

We hope ξ_X to be the horizontal lift[8] of $X^\sharp$; by construction $d\pi\xi_X = X^\sharp$. We have

$$\begin{aligned}
\alpha_{rk_0a}(\xi_X) &= \big[dL_{(rk_0a)^{-1}}\xi_X(rk_0a)\big]_{\mathcal{A}}\\
&= \frac{d}{dt}\Big[a^{-1}k_0^{-1}r^{-1}re^{tX}k_0a\Big]^{\mathcal{A}}_{t=0}\\
&= \frac{d}{dt}\Big[a^{-1}\,\mathbf{Ad}(k_0^{-1})e^{tX}a\Big]^{\mathcal{A}}_{t=0}\\
&= \big[\operatorname{Ad}(a^{-1}k_0^{-1})X\big]_{\mathcal{A}}.
\end{aligned}$$

One can, by brute force computation[9], show that the difference $\operatorname{Ad}(ak_0)X - \operatorname{Ad}(k_0)X$ is skew-diagonal when

$$X = \begin{pmatrix} a' & n\\ 0 & -a' \end{pmatrix}, \quad k_0 = \begin{pmatrix} \cos a & \sin a\\ \sin a & \cos a \end{pmatrix}, \quad a = \begin{pmatrix} a & 0\\ 0 & 1/p \end{pmatrix}.$$

So $\operatorname{Ad}(a)$ does not change the $\mathcal{A}$-component of $\operatorname{Ad}(k_0)X$. We conclude that

$$\alpha(\xi_X) = \big[\operatorname{Ad}(k_0^{-1})X\big]_{\mathcal{A}}. \tag{5.173}$$

When $X \in \mathcal{R}$, we consider $\tilde{X}_g = (dL_g)_eX$;

$$\tilde{X}_{rk_0a} = dL_{rk_0a}X, \tag{5.174}$$

in particular, $\tilde{X}_r = \frac{d}{dt}\Big[re^{tX}\Big]_{t=0}$. We denote by τ the action

$$\begin{aligned}
\tau_g \colon \mathcal{O} &\to \mathcal{O}\\
\tau_g \operatorname{Ad}(r)H &= \operatorname{Ad}(gr)H
\end{aligned} \tag{5.175}$$

In particular

$$\begin{aligned}
d\pi_g dL_g Y &= \frac{d}{dt}\Big[\pi(ge^{tY})\Big]_{t=0}\\
&= \frac{d}{dt}\Big[\operatorname{ad}(ge^{tY}H)\Big]_{t=0}\\
&= \frac{d}{dt}\Big[\tau_g \operatorname{Ad}(e^{tY})H\Big]_{t=0}\\
&= (d\tau_g)_H d\pi_e Y,
\end{aligned}$$

thus

$$d\pi \circ dL = d\tau \circ d\pi. \tag{5.176}$$

With definition (5.174), we have $\alpha(\tilde{X}) = X_{\mathcal{A}}$ because

$$\alpha\big(dL_{rk_0a}X\big) = \big[dL_{(rk_0a)^{-1}}dL_{rk_0a}X\big]_{\mathcal{A}} = X_{\mathcal{A}}.$$

We are now able to find some horizontal lift.

Proposition 5.54.
The horizontal lift of $X^\sharp$ is

$$\overline{X^\sharp} = \xi_X - [\operatorname{Ad}\widetilde{(k_0^{-1})}X]_{\mathcal{A}}.$$

[8] We will see in proposition 5.54 that it is not the case, but for the moment, we hope it.
[9] Or by remarking that $\mathcal{A}$ is abelian.

Proof. First, we have

$$\begin{aligned} d\pi\overline{X^\sharp}|_{rk_0a} &= d\pi\xi_X - d\pi(dL_{rk_0a})_e[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}} \\ &= \frac{d}{dt}\Big[\mathrm{Ad}(re^{tX}o)\Big]_{t=0} - d\tau\, d\pi[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}. \end{aligned}$$

The first term is $X^\sharp$ while the second is zero because if $A\in\mathcal{A}$,

$$\begin{aligned} d\pi A &= \frac{d}{dt}\Big[\pi(e^{tA})\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\mathrm{Ad}(e^{tA})H\Big]_{t=0} \\ &= 0. \end{aligned}$$

On the other hand,

$$\alpha(\overline{X^\sharp}) = \alpha(\xi_X) - [\mathrm{Ad}(k_0)^{-1}X]_{\mathcal{A}} = 0.$$

□

Now we prove that the function $\overline{X^\sharp}\cdot\hat\psi$ is equivariant, and therefore that the definition

$$\widehat{\nabla_{X^\sharp}\psi} = \overline{X^\sharp}\cdot\hat\psi$$

works. Using equivariance of $\hat\psi$,

$$\begin{aligned} \overline{X^\sharp}\cdot\hat\psi(ga_1) &= \frac{d}{dt}\Big[\hat\psi(\xi_X(t))\Big]_{t=0} - \hat\psi\big(dL_{ga_1}[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}\big) \\ &= \frac{d}{dt}\Big[\hat\psi\big(re^{tX}k_0aa_1\big)\Big]_{t=0} - \frac{d}{dt}\Big[\hat\psi\big(ga_1e^{t[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}}\big)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\rho(a_1)\hat\psi\big(re^{tX}k_0a\big)\Big]_{t=0} - \frac{d}{dt}\Big[\rho(a_1)\hat\psi\big(ge^{t[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}}\big)\Big]_{t=0}, \\ &= \rho(a_1)\hat\psi(\xi_X) - \frac{d}{dt}\Big[\rho(a_1)\hat\psi\big([\mathrm{Ad}\widetilde{(k_0^{-1})}X]_{\mathcal{A}}|_g\big)\Big]_{t=0} \\ &= \rho(a_1)\hat\psi(\xi_X) - \rho(a_1)[\mathrm{Ad}\widetilde{(k_0^{-1})}X]_{\mathcal{A}}\hat\psi(g) \\ &= \rho(a_1)(\overline{X^\sharp}\cdot\hat\psi)(g). \end{aligned} \tag{5.177}$$

for the third line, we used the fact that $\mathcal{A}$ is abelian

5.10.9.1 Clifford algebra for AdS_2

Our basis of $\mathcal{A}\oplus\mathcal{N}$ is

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad\text{and}\quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and we choose

$$o = \mathrm{Ad}(k_0)H = \cos(2k_0)H + \sin(2k_0)(E+F).$$

Since (at first order in t) $\mathrm{Ad}(e^{tH})o = \cos(2k_0)\mathbb{1} + \sin(2k_0)\big(E + 2tE + F - 2tF\big)$,

$$H^\sharp_{rk_0a} = 2\sin(2k_0)\,\mathrm{Ad}(r)(E-F),$$

and

$$E^\sharp_{rk_0a} = \mathrm{Ad}(r)\big(-2\cos(2k_0)E + \sin(2k_0)H\big).$$

We have to compute the metric matrix for this basis; we know from equation (3.24), the Killing form is Ad-invariant and $(\mathfrak{sl}(2,\mathbb{R}),B)\simeq(\mathbb{R}^3,\eta_{21})$. So the $\mathrm{Ad}(r)$ disappears in the computation of $B(X^\sharp,Y^\sharp)$. We get

$$\begin{aligned} B(H^\sharp,H^\sharp) &= 4\sin^2(2k_0)B(E-F,E-F) \\ &= -32\sin^2(2k_0) \\ B(E^\sharp,E^\sharp) &= \sin^2(2k_0)B(H,H) \\ &= 8\sin^2(2k_0) \\ B(E^\sharp,H^\sharp) &= -4\sin(2k_0)\cos(2k_0)B(E,E-F) + 2\sin^2(2k_0)B(H,E-F) \\ &= 16\sin^2(2k_0)\cos(2k_0). \end{aligned}$$

So the metric is in the basis $\{H^\sharp, E^\sharp\}$

$$g = \begin{pmatrix} -32\sin^2(2k_0) & 16\sin(2k_0)\cos(2k_0) \\ 16\sin(2k_0)\cos(2k_0) & 8\sin^2(2k_0) \end{pmatrix}. \tag{5.178}$$

When we consider the orbit of $E+F$, we choose $o = E+F$, i.e. $\cos(2k_0) = 0$, $\sin(2k_0) = 1$ so that

$$H^\sharp_{rk_0a} = 2\,\mathrm{Ad}(r)(E-F), \quad E^\sharp_{rk_0a} = \mathrm{Ad}(r)H, \tag{5.179}$$

and

$$g = \begin{pmatrix} -32 & 0 \\ 0 & 8 \end{pmatrix};$$

in the case of the orbit of $-(E+F)$, we get the same. The negative vector is $H^\sharp$ and the positive one is $E^\sharp$.

5.10.9.2 Identification $\mathcal{Q} \leftrightarrow \Lambda W$

We want a linear bijection $\phi\colon \mathcal{Q} \to \Lambda W$ such that

$$\rho(s)\phi(X) = \phi\big(\rho(s)X\big)$$

where the left hand side action of Spin is the usual on ΛW while the right hand side one remains to be defined. The implementation of this is easy: we can take any bijection between $\mathcal{Q}$ and ΛW and define

$$\rho(s)X = \phi^{-1}\big(\rho(s)\phi(X)\big). \tag{5.180}$$

Spinors on AdS_2 are given by equivariant functions $\hat{\psi}\colon \mathcal{U}_G \to \Lambda W$ which are now replaced by $\tilde{\psi}\colon R \to \mathcal{Q} \simeq \Lambda W$ by

$$\hat{\psi}(rk_0a) = \rho(a^{-1})\tilde{\psi}(r).$$

So the set of sections of the spinor bundle over $\mathcal{U}$ is

$$\Gamma_{\mathcal{U}} \simeq C^\infty(R, \Lambda W).$$

5.10.9.3 Covariant derivative

The aim is now to compute

$$\begin{aligned} \widetilde{\nabla_{X^\sharp}\psi}(r) &= \widehat{\nabla_{X^\sharp}\psi}(rk_0) \\ &= \overline{X^\sharp} \cdot \hat{\psi}|_{rk_0} \\ &= \big(\xi_X - [\mathrm{Ad}\widetilde{(k_0^{-1})}X]_{\mathcal{A}}\big) \cdot \hat{\psi}|_{rk_0} \\ &= \frac{d}{dt}\Big[\tilde{\psi}(re^{tX})\Big]_{t=0} - \frac{d}{dt}\Big[\rho\big(e^{t[\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}}\big)\Big]_{t=0}\tilde{\psi}(r) \\ &= \tilde{X}_r\tilde{\psi}(r) - d\rho_e\big([\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}\big)\tilde{\psi}(r). \end{aligned}$$

Our final formula for the covariant derivative is

$$\widetilde{\nabla_{X^\sharp}\psi}(r) = \tilde{X}_r\tilde{\psi} - d\rho\big([\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}\big)\tilde{\psi}. \tag{5.181}$$

The Dirac operator will be a linear combination of vectors of the form

$$\tilde{X} + d\rho\big([\mathrm{Ad}(k_0^{-1})X]_{\mathcal{A}}\big).$$

Notice that $\tilde{X}$ is left invariant and the second term is even independent of the point, so the whole is left invariant.

5.11 Dirac operator on AdS_3

The definition is

$$AdS_3 = \frac{\mathrm{SO}(2,2)}{\mathrm{SO}(1,2)},$$

and the group which acts is the AN of $\mathrm{SO}(2,2)$. The Lie algebra is given by

$$\begin{aligned} \mathcal{A} &= \{J_1, J_2\} \\ \mathcal{N} &= \{M, L\} \end{aligned}$$

which has dimension 4. So there is a stabiliser. One can prove that for the open orbit of $u = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, the stabiliser is $\{e^{aJ_2}\}$, i.e.

$$[e^{aJ_2}u] = [u]. \tag{5.182}$$

For the spin group, we find

$$\mathrm{Spin}(2,1) \simeq SL_2^*(\mathbb{R}),$$

the group of 2×2 matrices with determinant equals to ± 1 (cf [37]). Let us recall that the isomorphism $AdS_3 \simeq SL(2,\mathbb{R})$ is given by

$$SL(2,\mathbb{R}) = \begin{pmatrix} t+x & y-u \\ y+u & t-x \end{pmatrix}$$

with $u^2+t^2-x^2-y^2 = 1$. For sake of simplicity, we denote $SL(2,\mathbb{R})$ by G. It is explained in [49] that the map

$$\begin{aligned} \psi\colon (G \times G) \times AdS_3 &\to AdS_3 \\ (g_1,g_2)x &= g_1 x g_2^{-1} \end{aligned} \tag{5.183}$$

provides a local isomorphism $G \times G \simeq O(2,2)$. Moreover we have locally :

$$\frac{G \times G}{\mathbb{Z}_2} \simeq \mathrm{SO}(2,2).$$

At each point $x \in AdS_3$, we have an isomorphism

$$\mathrm{SO}(2,2)_x \simeq \mathrm{SO}(2,1)$$

where $\mathrm{SO}(2,2)_x$ is the stabiliser of x in $\mathrm{SO}(2,2)$. So we define the isomorphism

$$\chi_x\colon \mathrm{Spin}(2,1) \to \mathrm{SO}(2,2)_x$$

which is a double covering. If $d\psi\colon \mathcal{G} \oplus \mathcal{G} \to \mathfrak{so}(2,2)$ is the isomorphism of [49], we define $\psi\colon G \times G \to \mathrm{SO}(2,2)$ by

$$\psi(e^X) = e^{d\psi X},$$

which is a good definition because the exponential is surjective on $G \times G$. For each $x \in AdS_3$, we consider the isomorphism

$$\phi_x\colon \mathrm{SO}(2,1) \to \mathrm{SO}(2,2)_x$$

such that $\phi_x\big(\mathrm{SO}(2,1)\big) = \mathrm{SO}(2,2)_x$.

We define $\chi(s)_i\colon \mathrm{Spin}(2,1) \to \mathrm{SO}(2,2)$ by

$$\chi(s) = \chi(s)_1 v \chi(s)_2.$$

The choice of $\chi(s)_i$ is not unique. So we define the action of $\mathrm{Spin}(2,1)$ on $G \times G$ by

$$(g,h)\cdot s = \big(\chi(s)_1 g, \chi(s)_2^{-1} h\big). \tag{5.184}$$

Therefore we have

$$\begin{aligned} \psi\big((g,h)\cdot s\big)x &= \chi(s)_1 g x h^{-1} \chi(s)_2 \\ &= \chi(s)\big(gxh^{-1}\big) \\ &= \chi(s)\big(\psi(g,h)x\big). \end{aligned}$$

5.11.1 Spin structure on AdS_3

From previous considerations, the first choice should be

$$P = \frac{AN}{S} \times \mathrm{Spin}(2,1),$$

but it is easy to remark that $\mathcal{R}' = \{J_1, M, L\}$ is a Lie algebra. So we use the corresponding Lie group R instead of the homogeneous space AN/S (these two are isomorphic). Thus the choice is

$$P = R' \times \mathrm{Spin}(2,1), \tag{5.185}$$

with the projection $\pi\colon P \to AdS_3$,

$$\pi\big(r', s\big) = \left[r' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right]$$

We consider

$$\begin{aligned} \theta\colon R' &\to \mathcal{U} = Ro \\ r' &\mapsto ro = \left[r' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right]. \end{aligned} \tag{5.186}$$

The projection $\pi\colon P \to \mathcal{U}$ reads $\pi = \theta \circ \mathtt{pr}_1$,

$$\pi\big(r', s\big) = [ro].$$

This definition works because for all a, there exists a $h \in H$ such that

$$e^{aJ_2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} h,$$

from construction of S. Then we look at the following :

$$\begin{tikzcd} \mathrm{Spin}(2,1) \arrow[r, rightsquigarrow] & R' \times \mathrm{Spin}(2,1) \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & \mathrm{SO}\,(\mathcal{U}) \arrow[dl] & \mathrm{SO}(2,1) \arrow[l, rightsquigarrow] \\ & & \mathcal{U} & & \end{tikzcd}$$

The action of $\mathrm{Spin}(2,1)$ on P is

$$\big(r', s'\big) \cdot s = \big(r', s's\big).$$

In order to define φ, we consider the isomorphism

$$\phi_x\colon\ \mathrm{SO}(2,1) \to \mathrm{SO}(2,2)_x$$

between $\mathrm{SO}(2,1)$ and the stabiliser of x in $\mathrm{SO}(2,2)$. This extends to an automorphism

$$\phi_x\colon\ \mathrm{SO}(2,2) \to \mathrm{SO}(2,2),$$

Problem and misunderstanding 27.
I'm not sure of that extension, but we do not use it here.

and we define the action of $\mathrm{SO}(2,1)$ on $\mathrm{SO}\,\big(AdS_3\big)$ by

$$\{b_i\}_x \cdot g = \{\phi_x(g) b_i\}_x. \tag{5.187}$$

Then we define

$$\varphi\big(r', s\big) = \{\phi_{\pi r'}\big(\chi(s)\big) b_i\}_{\pi r'} \tag{5.188}$$

if $\{b_i\}$ is a reference basis at $\pi[r]$. So this construction implies the choice of a section of $\mathrm{SO}\,\big(AdS_3\big)$. Now, using the fact that both ϕ_x and χ are morphisms, we find

$$\begin{aligned} \varphi\big((r', s) \cdot s'\big) &= \big\{\phi_{\pi r'}\big(\chi(ss')\big)\big\}_{\pi r'} \\ &= \big\{\phi_{\pi r'}\big(\chi(s)\big) b_i\big\} \cdot \chi(s) \\ &= \varphi(r', s) \cdot \chi(s'). \end{aligned} \tag{5.189}$$

This proves that the construction gives a spin structure.

5.11.2 Connection on the spinor bundle

A left invariant vector on $\mathcal{U}$ is of the form

$$\tilde{X}_{xo} = \frac{d}{dt}\Big[x e^{tX} o \Big]_{t=0} = \frac{d}{dt}\Big[\pi\big(x e^{tX}, s\big) \Big]_{t=0}$$

for any $s \in \mathrm{Spin}(2,1)$. On AdS_3 (in fact on $\mathcal{U}$) we consider the left invariant vector field

$$X^{\sharp}_{[x]} = \frac{d}{dt}\Big[\big[x e^{tX}\big] \Big]_{t=0} \tag{5.190}$$

which leads us to consider the following field on P :

$$\xi_X(r', s) = \frac{d}{dt}\Big[r' e^{tX}, s\Big]_{t=0} \in T_{(r',s)}P. \tag{5.191}$$

This defines a field which projects to the left invariant field on $\mathcal{U}$:

$$d\pi\xi_X(r', s) = X^\sharp_{r'}. \tag{5.192}$$

Lemma 5.55.
On the general vector

$$\Sigma = \frac{d}{dt}\Big[r'(t), s(t)\Big]_{t=0}, \tag{5.193}$$

the formula

$$\alpha_{(r',s_0)}\Sigma = -\frac{d}{dt}\Big[s_0^{-1}s(t)\Big]_{t=0} \in \mathfrak{spin}(2,1) \tag{5.194}$$

where $s_0 = s(0)$ defines a connection form.

Proof. First let $A \in \mathfrak{spin}(2,1)$ and

$$A^*_\xi = \frac{d}{dt}\Big[\xi \cdot e^{-tA}\Big]_{t=0}.$$

We have

$$\alpha\big(A^*_{(r'),s_0}\big) = \alpha\frac{d}{dt}\Big[(r', s_0)\cdot e^{-tA}\Big]_{t=0} = \alpha\frac{d}{dt}\Big[(r', s_0 e^{-tA})\Big]_{t=0} = -\frac{d}{dt}\Big[s_0^{-1}s_0 e^{-tA}\Big]_{t=0} = A.$$

Now we take back the vector Σ of equation (5.193), an element $a \in \mathrm{Spin}(2,1)$ and we compute

$$\begin{aligned}
(dR_a\alpha)\Sigma &= \alpha\frac{d}{dt}\Big[\big(r'(t), s(t)\big)\cdot a\Big]_{t=0}\\
&= \alpha\frac{d}{dt}\Big[\big(r'(t), s(t)a\big)\Big]_{t=0}\\
&= -\frac{d}{dt}\Big[a^{-1}s(0)^{-1}s(t)a\Big]_{t=0}\\
&= -\operatorname{Ad}(a^{-1})\frac{d}{dt}\Big[s(0)^{-1}s(t)\Big]_{t=0}\\
&= \operatorname{Ad}(a^{-1})\alpha(\Sigma).
\end{aligned}$$

□

Thus that is a connection. This is however not the spin connection. Let β be the Levi-Civita connection on the frame bundle $\mathrm{SO}(AdS_3)$. If

$$\Sigma = \frac{d}{dt}\Big[r'(t), s(t)\Big]_{t=0},$$

we have

$$\beta d\phi\Sigma = \phi_{r'}\big(\chi(s_0)\big)^{-1}\frac{d}{dt}\Big[\phi_{r'(t)}\big(\chi(s_t)\big)\Big]_{t=0}\Big|_{\mathcal{H}}. \tag{5.195}$$

If we note $\phi_{r'(t)}\big(\chi(s_t)\big) = \phi\big(r'(t), \chi(s_f)\big)$, the derivative in (5.195) with respect to t reads

$$\frac{d}{dt}\Big[\phi\big(r'(t), \chi(s_0)\big)\Big]_{t=0} + \frac{d}{dt}\Big[\phi\big(r', \chi(s_t)\big)\Big]_{t=0}. \tag{5.196}$$

The second term of $\beta d\varphi\Sigma$ is

$$\begin{aligned}
\frac{d}{dt}\Big[\phi_{r'}(\chi(s_0))^{-1}\phi_{r'}(\chi(s_t))\Big]_{t=0}\Big|_{\mathcal{H}} &= \frac{d}{dt}\Big[\phi_{r'}\big(\chi(s_0^{-1}s_t)\big)\Big]_{t=0}\Big|_{\mathcal{H}}\\
&= d\phi_{r'}d\chi(s_0^{-1}s'(0))\big|_{\mathcal{H}}.
\end{aligned}$$

From all that we want to define

$$\alpha^S_{(r',s_0)}\Sigma = d\phi d\chi\big(s_0^{-1}s'(0)\big)\big|_{\mathcal{H}} + \phi_{r'}\big(\chi(s_0)\big)^{-1}\frac{d}{dt}\Big[\phi_{r'(t)}\chi(s_0)\Big]_{t=0}\Big|_{\mathcal{H}}, \tag{5.197}$$

and we would not have $\alpha^S(\xi_X) = 0$.

5.11.3 Horizontal lift

Since the spin component of the path of ξ_X is constant, we have $\alpha(\xi_X) = 0$, so equation (5.192) says that

$$\overline{X^\sharp} = \xi_X. \tag{5.198}$$

Let us recall that an equivariant function (which defined a section of an associated bundle) is

$$\begin{aligned} \hat{\psi} &\colon P \to V \\ \hat{\psi}(\xi \cdot g) &= \rho(g^{-1})\hat{\psi}(\xi). \end{aligned} \tag{5.199}$$

General definition of an equivariant derivative (theorem 4.25) leads to

$$\widehat{\nabla_{X^\sharp}\psi} = \overline{X^\sharp} \cdot \hat{\psi} = \xi_X \cdot \hat{\psi}.$$

In our setting, the equivariance of $\hat{\psi}$ reads, for all $a \in \mathrm{Spin}(2,1)$,

$$\hat{\psi}\big(([r], s) \cdot a\big) = \hat{\psi}\big([r], sa\big) \stackrel{!}{=} \rho(a^{-1})\hat{\psi}\big([r], s\big).$$

We check the equivariance of $\widehat{\nabla_{X^\sharp}\psi}$ by the following computation :

$$\begin{aligned} \widehat{\nabla_{X^\sharp}\psi}\big(([r], s) \cdot a\big) &= \widehat{\nabla_{X^\sharp}\psi}([r], sa) \\ &= (\xi_X \cdot \hat{\psi})([r], sa) \\ &= \frac{d}{dt}\Big[\hat{\psi}\big([re^{tX}], sa\big)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\rho(a^{-1})\hat{\psi}\big([re^{tX}], s\big)\Big]_{t=0} \\ &= \rho(a^{-1})(\xi_X \cdot \hat{\psi})\big([r], s\big) \\ &= \rho(a^{-1})\widehat{\nabla_{X^\sharp}\psi}\big([r], s\big). \end{aligned}$$

We define $\tilde{\psi}\colon AN/S \to \Lambda W$ by

$$\tilde{\psi}([r]) = \hat{\psi}([r], e),$$

so that

$$\hat{\psi}([r], s) = \rho(s^{-1})\tilde{\psi}([r]). \tag{5.200}$$

We can conclude

$$\begin{aligned} \widetilde{\nabla_{X^\sharp}\psi}([r]) &= \widehat{\nabla_{X^\sharp}\psi}([r], e) \\ &= \xi_X\hat{\psi}([r], e) \\ &= \frac{d}{dt}\Big[\hat{\psi}\big([re^{tX}], e\big)\Big]_{t=0} \\ &= \frac{d}{dt}\Big[\tilde{\psi}\big([re^{tX}]\big)\Big]_{t=0} \\ &= \tilde{X}_{[r]}\tilde{\psi}([r]). \end{aligned}$$

So

$$\widetilde{\nabla_{X^\sharp}\psi} = \tilde{X}_{[r]}\tilde{\psi}. \tag{5.201}$$

5.11.4 Spin structure on AdS_3

5.11.4.1 Spin structure on the whole AdS_3

Problem and misunderstanding 28.
The following seems to contradict what I find in Michelson-Donaldson

The central fact is that

$$\mathrm{Spin}(2,1) \simeq \Delta \simeq \mathrm{SL}(2,\mathbb{R})$$

where $\Delta = \{(g,g) \text{ st } g \in \mathrm{SL}(2,\mathbb{R})\} \subset G_0$. We take as notations: $G_0 = \mathrm{SL}(2,\mathbb{R})$ and $\overline{G} = G_0 \times G_0$.

Lemma 5.56.
We have the following homogeneous space isomorphism :

$$\overline{G}/\Delta \simeq \mathrm{SL}(2,\mathbb{R}).$$

Proof. We have an action $\overline{G} \times AdS_3 \to AdS_3$,

$$(g,h)x = gxh^{-1} \tag{5.202}$$

where $x \in \mathrm{SL}(2,\mathbb{R})$ is seen as in AdS_3 by the usual isomorphism. Moreover we consider the isomorphism

$$\overline{G}/\Delta \simeq \mathrm{SL}(2,\mathbb{R}) \tag{5.203}$$
$$[x_1, x_2] \mapsto x_1 x_2^{-1} \tag{5.204}$$

which is well defined because $[x_1 g, x_2 g] \mapsto x_1 g g^{-1} x_2^{-1} = x_1 x_2^{-1}$. In particular, $[g,g] \mapsto e \in \mathrm{SL}(2,\mathbb{R})$. So $\overline{G}$ acts on $\mathrm{SL}(2,\mathbb{R})$ and the elements which fix e are the one of Δ. It proves the lemma. □

We are going to take the following structure :

$$\begin{tikzcd} \overline{G} \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & \overline{G}/\mathbb{Z}_2 \arrow[dl] \\ & M & \end{tikzcd} \tag{5.205}$$

where M is G_0 seen as $M = \overline{G}/\Delta \simeq \mathrm{SL}(2,\mathbb{R}) \simeq AdS_3$, and the projection $\pi\colon \overline{G} \to M$ is given by $\pi(g,h) = gh^{-1}$. The action of $\Delta \simeq \mathrm{Spin}(2,1)$ on $\overline{G}$ is given by formula $(xg,g)\cdot(a,a) = (xga,ga)$. First, let us prove the following.

Proposition 5.57.
The frame bundle over AdS_3 can be seen as

$$\mathrm{SO}(AdS_3) \simeq \overline{G}/\mathbb{Z}_2$$

where $\overline{G} = \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{R})$.

Proof. In the fiber bundle $\pi\colon \overline{G} \to M$, the fibre over $x \in \mathrm{SL}(2,\mathbb{R})$ is the set of (g,h) such that $gh^{-1} = x$, or

$$\overline{G}_x = \{(xg,g)\} \subset \overline{G}.$$

We will give a surjective map $\overline{G}_x \to \mathrm{SO}(M)_x$, the fibre of the frame bundle over $x \in AdS_3$. For this, we see a basis of AdS_3 as an isometric map $b\colon \mathcal{G}_0 \to T_xM$ where $\mathcal{G}_0 = \mathfrak{sl}(2,\mathbb{R})$, and we define

$$\begin{aligned} \psi_x\colon \overline{G}_x &\to \mathrm{SO}(M)_x \\ \psi_x(xg,g)(X) &= (dL_x)_e\big(\mathrm{Ad}(g^{-1})X\big) \end{aligned} \tag{5.206}$$

for all $X \in \mathcal{G}_0$. Let us study the kernel of this map, i.e. elements such that $\psi(xg_1,g_1) = \psi(xg_2,g_2)$. It needs, for all $X \in \mathfrak{sl}(2,\mathbb{R})$,

$$\mathrm{Ad}(g_1^{-1})X = \mathrm{Ad}(g_2^{-1})X,$$

but we know that the requirement $\mathrm{Ad}(g)X = X$ is the fact the g is in the center of the group. In our case, it results that $g_2^{-1}g_1 = \pm\,\mathrm{id}$, so

$$\psi(xg_1,g_1) = \psi(\pm xg_1, \pm g_1)$$

where the same $\pm$ has to be taken in both appearances of the right hand side. Now we put all the ψ_x together to get $\psi\colon \overline{G} \to \mathrm{SO}(M)$. Once again we look in which cases $\psi(g_1,h_1) = \psi(g_2,h_2)$. We put this condition under the form

$$\psi(g_1h_1^{-1}h_1, h_1) = \psi(g_2h_2^{-1}h_2, h_2)$$

which immediately gives $h_1 = \pm h_2$. But on the other hand the base point of $\psi(g_ih_i^{-1}, h_i)$ is $g_ih_i^{-1}$, so that the condition also ask $g_1h_1^{-1} = g_2h_2^{-1}$ which in turn gives $g_1 = \pm g_2$ with the same $\pm$ as in $h_1 = \pm h_2$. We conclude that $\mathbb{Z}_2$ is the problem for the inverse of ψ. This proves the proposition. □

We will usually use the same notation, ψ, to denote the map from $\overline{G}$ and the one from $\overline{G}/\mathbb{Z}_2$. The following lemma will prove useful to study the actions of the structure groups in the picture (5.205).

Lemma 5.58.
The map

$$\begin{aligned} \mathrm{SL}(2,\mathbb{R}) &\to \mathrm{SO}_0(1,2) \\ g &\mapsto \mathrm{Ad}(g). \end{aligned} \tag{5.207}$$

is a double covering.

Proof. No proof. □

The action of $a \in \mathrm{SO}_0(1,2)$ on $(xg,g) \in \overline{G}/\mathbb{Z}_2$ is defined by

$$\psi\big((xg,g)\cdot a\big) = (dL_x)_e \,\mathrm{Ad}(a^{-1}g^{-1}). \tag{5.208}$$

On the other hand, let us see how does $(a,a) \in \Delta \simeq \mathrm{Spin}(2,1)$ acts on $\overline{G}$ and how does it reflects on the ψ level. Since $(xg,g)\cdot(a,a) = (xga,ga)$, we have

$$\psi\big([xg,g]\cdot a\big) = \psi\big((xg,g)\cdot(a,a)\big),$$

and then

$$\varphi\big((xg,g)\cdot a\big) = (xg,g)\cdot(a,a).$$

This proves that our structure is a spin structure.

5.11.4.2 Reduction to one open orbit

We will use this isomorphism between AdS_3 and $\mathrm{SL}(2,\mathbb{R})$:

$$\begin{pmatrix} u \\ t \\ x \\ y \end{pmatrix} \mapsto \begin{pmatrix} u+x & y-t \\ y+t & u-x \end{pmatrix}.$$

Then the famous point $[u] = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in AdS_3$ corresponds to the element $J := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$. This is our base point of the open orbit. We could also take

$$k_0 = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \in K_0$$

where K_0 is the "K" of $\mathrm{SL}(2,\mathbb{R})$.

Problem and misunderstanding 29.
I think that J is also a complex structure. To be checked.

We have $J = k_0^2$ and following the action (5.202), we have $J = (k_0, k_0^{-1})e$. The subgroup $\overline{R} \subset \overline{G}$ acts on AdS_3, and we want to know the stabilizer of J. The condition is $(r,r')\cdot J = J$, or

$$r = \mathbf{Ad}(J)r',$$

but $\mathbf{Ad}(J) = \theta$ (the Cartan involution). So an element $(r,r') \in \overline{R}$ stabilises J if it is of the form $(r,\theta r)$, thus

$$\mathfrak{s} = \text{Lie algebra of the stabiliser of } J = \{(X,\theta X) \text{ st } X \in \mathcal{R}_0\} \cap \mathcal{R},$$

where the intersection with $\mathcal{R}$ is important because θ can send out of $\mathcal{R}_0$. Note that when X has a $\mathcal{N}$ component, then θX has a $\overline{\mathcal{N}}$ component, so $(X,\theta X) \in (\mathcal{A}\oplus\mathcal{N}, -\mathcal{A}\oplus\overline{\mathcal{N}})$ where the minus sign comes from the fact that $\theta(\mathcal{A}) = -\mathcal{A}$. Then X cannot have a $\mathcal{N}$ component and finally,

$$\mathfrak{s} = \mathbb{R}(H,-H) \in \mathcal{Q}.$$

The group R' is

$$R' = e^{\mathcal{R}'} = \{(an, an') \text{ st } n,n' \in N_0\} \tag{5.209}$$

because $\mathcal{R}'$ is $\mathcal{R}$ minus the stabiliser, i.e. $\mathcal{R}' = \mathbb{R}(H,H)\oplus\mathcal{N}$. We have the identification $r' \mapsto r'\cdot J$ between R' and the open orbit $\mathcal{U}$. As usual, the action is $(g,h)\cdot x = gxh^{-1}$ if $r' = (g,h)$. Notice in particular that $R' \neq R_0' \times R_0'$.

Up to now we studied the fiber $\overline{G} \to M$; we are now able to restrict it to $\overline{G}|_{\mathcal{U}} \to \mathcal{U}$ and to establish an isomorphism with the trivial bundle $R' \times G_0 \to R'$. The fiber over $x \in \mathcal{U}$ is

$$\overline{G}_x = \{(xg, g)\}.$$

We define the isomorphism as follows:

$$\begin{aligned} \tau\colon R' \times G_0 &\to \overline{G}|_{\mathcal{U}} \\ (r', g) &\mapsto (r' \cdot Jg, g) \end{aligned} \tag{5.210}$$

and we have the following picture:

$$\begin{tikzcd} R' \times G_0 \arrow[r, "\tau"] \arrow[d] & \overline{G}|_{\mathcal{U}} \arrow[d, "\pi"] \\ R' \arrow[r, dotted, "\tau"] & \mathcal{U} \end{tikzcd}$$

in which are defines by

$$\begin{tikzcd} (r', g) \arrow[r, "\tau"] \arrow[d] & (r' \cdot Jg, g) \arrow[d, "\pi"] \\ r' \arrow[r, dotted, "\tau"] & r' \cdot J \end{tikzcd}$$

where the dotted line denotes the induced map from τ, which is denoted by the same symbol. The map $\tau\colon R' \to \mathcal{U}$ is just the restriction of the original τ to $g = e$. Notice that this τ provides a diffeomorphism of the basis spaces R' and $\mathcal{U}$.

5.11.4.3 Spin connection

The spin connection on $\overline{G}|_{\mathcal{U}}$ is given by

$$\alpha^S_{(g,h)}\Sigma = \big[dL_{(g,h)^{-1}}\Sigma\big]_{\mathcal{H}}, \tag{5.211}$$

or

$$\alpha^S_{(g,h)} = \mathtt{pr}_{\mathcal{H}} \circ \big(dL_{(g,h)^{-1}}\big)_{(g,h)}. \tag{5.212}$$

Notice that when we write $\mathcal{H}$, we think about Δ : the group by which quotient $\overline{G}$ in order to get $\mathrm{SL}(2,\mathbb{R}) \simeq AdS_3$. Our task now is to transfer this connection to $R' \times G_0$ by defining $\alpha' = \tau^*\alpha^S$. If $\Sigma \in T_{(r',g)}(R' \times G_0)$, we define

$$\alpha'_{(r',g)}\Sigma = \alpha^S(d\tau\Sigma). \tag{5.213}$$

Let us take $X \in \mathcal{G}_0$ and $0 \in \mathcal{R}'$ and let us compute $d\tau(0, X)$. More precisely, we consider

$$\begin{aligned} d\tau(0 \oplus \tilde{X}_g)_{(r',g)} &= d\tau\frac{d}{dt}\Big[r', ge^{tX}\Big]_{t=0} \\ &= \frac{d}{dt}\Big[r' \cdot Jge^{tX}, ge^{tX}\Big]_{t=0} \\ &= \big(\tilde{X}_{(r'\cdot Jg)}, \tilde{X}_g\big). \end{aligned}$$

The next step is to compute $d\tau\Sigma$ in the case where $\Sigma = (\underset{\sim}{Y} \oplus -1)_{(r',g)}$ with $Y \in \mathcal{R}' \subset \mathcal{R}_0 \oplus \mathcal{R}_0$. We have

$$d\tau\Sigma = \frac{d}{dt}\Big[\tau\big(e^{tY}r', g\big)\Big]_{t=0} \tag{5.214}$$

$$= \frac{d}{dt}\Big[(e^{tY}r' \cdot J)g, g\Big]_{t=0} \tag{5.215}$$

where, if $r' = (r_1, r_2)$, we consider $Y = \big((\underset{\sim}{Y}_1)_{r_1}, (\underset{\sim}{Y}_2)_{r_2}\big)$. This appears to be difficult to be computed. This reflects the fact that the connection should be complicated in the trivial bundle $R' \times G_0$.

But there are no fate. We remember that τ furnish a diffeomorphism between the basis spaces, so one can consider the bundle

$$\begin{tikzcd} \overline{G}|_{\mathcal{U}} \arrow[d, "\tau^{-1} \circ \pi"] \\ R' \end{tikzcd}$$

Vectors of $\mathcal{H}$ are of the form (X,X) with $X\in\mathfrak{sl}(2,\mathbb{R})$, thus $A\in T_{(xg,g)}\overline{G}|_{\mathcal{U}}$ fulfils $\alpha^S(A)=0$ if and only if

$$dL_{(xg,g)^{-1}}(A)=(X,-X)$$

for a certain $X\in\mathfrak{sl}(2,\mathbb{R})$. All this makes that the horizontal space over (xg,g) is given by

$$\operatorname{hor}(xg,g)=\big\{(\tilde{X}_{xg},-\tilde{X}_g)\text{ st }X\in\mathcal{G}_0=\mathfrak{sl}(2,\mathbb{R})\big\}. \tag{5.216}$$

The strategy now is to project that on R' and express Dirac operator in terms of the result. Let us make this simple computation :

$$\begin{aligned}
d\pi(\tilde{X}_{xg},\tilde{X}_g)&=\frac{d}{dt}\Big[\pi\big(xge^{tX},ge^{-tX}\big)\Big]_{t=0}\\
&=\frac{d}{dt}\Big[xge^{tX}e^{tX}g^{-1}\Big]_{t=0}\\
&=\frac{d}{dt}\Big[xe^{2t\operatorname{Ad}(g)X}\Big]_{t=0}\\
&=2(dL_x)_e\operatorname{Ad}(g)X.
\end{aligned}$$

This result has to be brought from $\mathcal{U}$ to R' by τ^{-1}. Now we take a $\tilde{Y}\in\mathfrak{X}(R')$ and we want to know which is the corresponding X, i.e. the $X\in\mathfrak{sl}(2,\mathbb{R})$ such that

$$d\tau^{-1}d\pi(\tilde{X}_{xg},-\tilde{X}_g)=\tilde{Y}.$$

From the previous computation, $\tilde{Y}=2d\tau^{-1}dL_x\operatorname{Ad}(g)X$, so

$$X=\frac{1}{2}\operatorname{Ad}(g^{-1})dL_{x^{-1}}d\tau\tilde{Y}. \tag{5.217}$$

We now precise our idea:

$$\tilde{Y}_{(r_1,r_2)}=\big((\tilde{Y}_1)_{r_1},(\tilde{Y}_2)_{r_2}\big)=\frac{d}{dt}\Big[r_1e^{tY_1},r_2e^{tY_2}\Big]_{t=0} \tag{5.218}$$

for $Y_i\in\mathcal{R}'_0$ and r_1, $r_2\in R_0$. In this case, the "x" in equation (5.217) is $(r'\cdot J)^{-1}$. Let us begin by taking $s'\in R'$ and compute $L_{(r'.J)^{-1}}\tau(s')$. Remember that $r'\cdot J=r_1Jr_2^{-1}$ from the general action (5.202), so if $r'=(r_1,r_2)$,

$$\begin{aligned}
dL_{(r'.J)^{-1}}\tau(s')&=(r'\cdot J)^{-1}s_1Js_2^{-1}\\
&=(r_1Jr_2^{-1})^{-1}s_1Js_2^{-1}\\
&=-r_2Jr_2^{-1}s_1Js_2^{-1}.
\end{aligned}$$

Now, we apply that result on computation of (5.217) with (5.218) :

$$\begin{aligned}
dL_{(r'.J)^{-1}}d\tau\tilde{Y}&=\frac{d}{dt}\Big[-r_2Jr_1^{-1}r_1e^{tY_1}Je^{-tY_2}r_2^{-1}\Big]_{t=0}\\
&=\frac{d}{dt}\Big[\mathbf{Ad}(r_2)\big(-Je^{tY_1}Je^{-tY_2}\big)\Big]_{t=0}\\
&=\frac{d}{dt}\Big[\mathbf{Ad}(r_2)e^{-tY_2}\Big]_{t=0}+\operatorname{Ad}(r_2)\operatorname{Ad}(J)Y_1\\
&=-\operatorname{Ad}(r_2)Y_2+\operatorname{Ad}(r_2)\theta(Y_1),
\end{aligned}$$

and finally,

$$\begin{aligned}
X&=\frac{1}{2}\operatorname{Ad}(g^{-1})dL_{(r'.J)^{-1}}d\tau\tilde{Y}\\
&=\frac{1}{2}\operatorname{Ad}(g^{-1})\big[\operatorname{Ad}(r_2)\theta(Y_1)-\operatorname{Ad}(r_2)Y_2\big].
\end{aligned} \tag{5.219}$$

For this X, the horizontal lift of $\tilde{Y}\in\mathfrak{X}(R')$ is $(X,-X)\in T\overline{G}|_{\mathcal{U}}$.

5.11.5 Left invariance of Dirac

Sections of the spin bundle over the open orbit $\mathcal{U}$ are given by equivariant functions $\hat{\psi}\colon\overline{G}|_{\mathcal{U}}\to\mathbb{R}^2$. The action of $\Delta\simeq\operatorname{Spin}(2,1)$ on $\overline{G}$ is

$$(g,h)\cdot(a,a)=(ga,ha).$$

We define $\tilde{\psi}$ by

$$\tilde{\psi}(g)=\hat{\psi}(g,e) \tag{5.220}$$

for $g \in \mathcal{U}$. We get back the original $\hat{\psi}$ by formula

$$\hat{\psi}(g,h) = \rho(h,h)^{-1}\tilde{\psi}(gh^{-1}). \tag{5.221}$$

Our intention is now to compute $\widehat{\nabla_Z\psi}(\xi) = \overline{Z}_\xi(\hat{\psi})$ with $\xi = (xg,g) \in \overline{G}|_{\mathcal{U}}$ (hence $x \in \mathcal{U}$) and $Z \in \mathfrak{X}(R')$. For instance we choose a left invariant $Z = \tilde{Y} = (\tilde{Y}_1, \tilde{Y}_2)$ for $Y_1, Y_2 \in \mathcal{R}_0'$. Recall that $\tilde{Y}$ is given by equation (5.218). From definition of the covariant derivative associated with the connection,

$$\widehat{\nabla_{\tilde{Y}}\psi}(\xi) = \overline{\tilde{Y}}_\xi(\hat{\psi}) = \overline{\tilde{Y}}_{(xg,g)}(\hat{\psi})$$

where $\overline{\tilde{Y}}_{(xg,g)}$ is an horizontal vector at (xg,g) whose projection is $\tilde{Y}$. From our previous work,

$$\overline{\tilde{Y}}_{xg,g} = (\tilde{X}_{xg}, -\tilde{X}_g)$$

with $X = \frac{1}{2}\operatorname{Ad}(g^{-1})\big(\operatorname{Ad}(r_2)\theta Y_1 - \operatorname{Ad}(r_2)Y_2\big)$. Let us understand the link between (r_1, r_2) and g, x. The vector $(\tilde{X}_{xg}, \tilde{X}_g)$ actually projects to a vector at $\tau^{-1}\circ\pi(xg,g) = \tau^{-1}(x)$. The fact that $x \in \mathcal{U}$ guarantees existence and uniqueness of $(r_1,r_2) \in R'$ such that $r_1 J r_2^{-1} = x$. We have

$$\begin{aligned}
\widetilde{\nabla_{\tilde{Y}}\psi}(x) &= \widehat{\nabla_{\tilde{Y}}\psi}(x,e) \\
&= (\tilde{X}_x, -\tilde{X}_e)\hat{\psi} \\
&= \frac{d}{dt}\hat{\psi}\Big(xe^{tX}, e^{-tX}\Big)\Big|_{t=0} \\
&= \frac{d}{dt}\rho(e^{tX}, e^{tX})\tilde{\psi}(xe^{2tX})\Big|_{t=0}.
\end{aligned}$$

The first term of the derivation (the one with $t = 0$ in the ρ) gives $2\tilde{X}_x\tilde{\psi}$. This is left invariant. The second is

$$\frac{d}{dt}\rho(e^{tX}, e^{tX})\tilde{\psi}(x)\Big|_{t=0}.$$

We want to test the condition (**??**) on this term. Let us pose

$$(E\tilde{\psi})(x) = (\tilde{X}_x, \tilde{X}_e)\hat{\psi} = \frac{d}{dt}\rho(e^{tX}, e^{tX})\tilde{\psi}(x)\Big|_{t=0}$$

with X given by equation (5.219). On the one hand,

$$L_y(E\tilde{\psi})(x) = (E\tilde{\psi})(yx) = \frac{d}{dt}\rho(e^{tX_a}, e^{tX_a})\tilde{\psi}(yx)\Big|_{t=0} \tag{5.222a}$$

with

$$X_a = \frac{1}{2}\Big(\operatorname{Ad}(r_2)\theta Y_1 - \operatorname{Ad}(r_2)\Big)Y_2 \tag{5.222b}$$

where (r_1, r_2) is given by yx. On the other hand,

$$E(L_y\tilde{\psi})(x) = \frac{d}{dt}\rho(e^{tX_b}, e^{tX_b})\tilde{\psi}(yx)\Big|_{t=0} \tag{5.223a}$$

with

$$X_b = \frac{1}{2}\big(\operatorname{Ad}(s_2)\theta Y_1 - \operatorname{Ad}(s_2)Y_2\big) \tag{5.223b}$$

where (s_1, s_2) is given by x.

The problem is that the choice of y is arbitrary, so that X_a and X_b could be too different. Ok. That's the proof that Dirac is not invariant. Here is the proof that Dirac is invariant.

Following equation (5.211), the spin connection form is

$$\alpha^S_{(g,h)}\Sigma = \big(dL_{(g,h)^{-1}}\Sigma\big)_{\mathcal{H}}.$$

If $L_{(x,y)}$ is the left translation by (x,y) we have

$$\big(L^*_{(x,y)}\alpha\big)_{(g,h)}\Sigma = \alpha_{(xg,yh)}\big(dL_{(x,y)}\Sigma\big) = \big(dL_{(g,h)^{-1}}\Sigma\big)_{\mathcal{H}}.$$

Thus we have $L^*_{(x,y)}\alpha^S = \alpha^S$. Now we consider the formula $\widehat{\nabla \cdot \psi}(\xi) = (\tilde{X}_{xg}, -\tilde{X}_g)\hat{\psi}$, and we will check that

$$(L_\eta \widehat{\nabla_Z\psi})(\xi) = \nabla_Z \widehat{(L_\eta\psi)}(\xi). \tag{5.224}$$

with $\xi = (xg, g)$ and $\eta = (a, b)$. On the one hand,

$$\begin{aligned}
(L_{(a,b)}\widehat{\nabla_Z\psi})(xg, g) &= \widehat{\nabla_Z\psi}(axg, bg)\\
&= \widehat{\nabla_Z\psi}(axgg^{-1}b^{-1}bg, bg)\\
&= (\tilde{X}_{(axb^{-1})bg}, -\tilde{X}_{bg})\hat{\psi}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\nabla_Z\widehat{(L_{(a,b)}\psi)}(xg, g) &= (\tilde{X}_{xg}, -\tilde{X}_g)\widehat{L_{(a,b)}\psi}\\
&= \frac{d}{dt}\widehat{L_{(a,b)}\psi}\big(xge^{tX}, ge^{-tX}\big)\Big|_{t=0}\\
&= \frac{d}{dt}\hat{\psi}\big(axge^{tX}, bge^{-tX}\big)\Big|_{t=0}\\
&= (\tilde{X}_{axg}, -\tilde{X}_{bg})\hat{\psi}\\
&= (\tilde{X}_{(axb^{-1})bg}, -\tilde{X}_{bg})\hat{\psi}.
\end{aligned}$$

5.12 Dirac operator on AdS_4

5.13 Dirac operator on AdS_l

5.13.1 Frame bundle

Construction of the frame bundle and the spin structure is a straightforward adaptation of theorem 2.2 (chapter ???) in [50], while Dirac operator and connection issues are adapted from proposition 1.3 (chapter III)

A **basis** of a m dimensional vector space V is a free and generating part; it only has the structure of a set. A frame of the vector space V is a nondegenerate map $b\colon \mathbb{R}^m \to V$. Let us give an example in three dimensions the difference. If $\{v_1, v_2, v_3\}$ is a basis of V, of course $\{v_2, v_1, v_3\}$ is the same basis. Order has no importance. But if $\{e_1, e_2, e_3\}$ is the canonical basis of $\mathbb{R}^3$, the *frames* $b(e_1) = v_1$, $b(e_2) = v_2$, $b(e_3) = v_3$ and $c(e_1) = v_2$, $c(e_2) = v_1$, $c(e_3) = v_3$ are not the same.

Now we consider $AdS_l = G/H = \mathrm{SO}(2, l-1)/\mathrm{SO}(1, l-1)$, the Lie algebra $\mathcal{G}$ has a reductive homogeneous space decomposition $\mathcal{G} = \mathcal{Q} \oplus \mathcal{H}$ and we consider the canonical projection $\pi\colon G \to AdS_l$.

Let the map (see relation (**??**))

$$\begin{aligned}
\alpha\colon H &\to \mathrm{SO}(\mathcal{Q})\\
h &\mapsto \mathrm{Ad}(h)|_{\mathcal{Q}}.
\end{aligned} \tag{5.225}$$

We consider, on $G \times \mathrm{SO}(\mathcal{Q})$, the equivalence relation $(g, A) \sim (g', A')$ if and only if there exists $h \in H$ such that $g' = gh$ and $A' = \alpha(h^{-1})A$. We denote by $G \times_\alpha \mathrm{SO}(\mathcal{Q})$ the set of equivalence classes. Now we have a principal bundle

$$\begin{tikzcd}
\mathrm{SO}(\mathcal{Q}) \arrow[r, squiggly] & G \times_\alpha \mathrm{SO}(\mathcal{Q}) \arrow[d, "p"]\\
& G/H
\end{tikzcd} \tag{5.226}$$

where $p[g, A] = [g]$ and the action is given by $[g, A]\cdot B = [g, A]$. The fact that the projection fulfils $p\big([g, A]\cdot B\big) = p[g, A]$ is evident, and the fact that the action is well defined is a simple computation : if $[g', A'] = [g, A]$, we have a $h \in H$ such that

$$[g', A'] \cdot B = [g', A'B] = [gh, \alpha(h^{-1})AB] = [g, AB] = [g, A] \cdot B.$$

Proposition 5.59.

Let $\tau(g)\colon AdS_l \to AdS_l$ be the action of $g \in G$ on AdS_l : $\tau(g)[g'] = [gg']$, and B be the frame bundle. We also consider the map $\sigma\colon \mathbb{R}^{1,l-1} \to \mathcal{Q}$ the isometry which sends the canonical basis of $\mathbb{R}^{1,l-1}$ to the usual basis $\{q_0, q_1, \ldots, q_{l-1}\}$ of $\mathcal{Q}$. The map

$$\begin{aligned}
\beta\colon G \times_\alpha \mathrm{SO}(\mathcal{Q}) &\to B\\
[g, A] &\mapsto d\tau(g)_\vartheta A \circ \sigma
\end{aligned} \tag{5.227}$$

provides a principal bundle isomorphism between the principal bundle (5.226) and the frame bundle over AdS_l.

By abuse of notation, we will not always write the σ.

Proof. We have to prove first that the map $\beta\colon G\times \mathrm{SO}(\mathcal{Q})\to B$ respects the classes. For that, consider $(g,A)\sim(g',A')$ and remark that

$$\begin{aligned}\beta(gh,\alpha(h^{-1})) &= d\tau(gh)_{\vartheta}\alpha(h^{-1})A = d\tau(g)d\tau(h)d\pi\,\mathrm{Ad}(h^{-1})d\pi^{-1}A\\ &= d\tau(g)d\tau(h)d\pi\,\mathrm{Ad}(h^{-1})d\pi^{-1}A = d\tau(g)d\pi dR_h d\pi^{-1}A\\ &= d\tau(g)_{\vartheta}A = \beta(g,A).\end{aligned}$$

where we used equation (**??**) and the fact that $\pi\circ L_g=\tau(g)\circ\pi$. The frame bundle is

$$\begin{tikzcd} \mathrm{SO}(1,l-1) \arrow[r, squiggly] & B \arrow[d, "p"] \\ & G/H \end{tikzcd} \tag{5.228}$$

where the fibre $B_{[g]}$ in B over $[g]$ is the set of isometric maps $\mathbb{R}^{1,l-1}\to T_{[g]}(AdS_l)$. So an element of B is of the form $\big([g],\tilde f\circ\sigma\big)$ where $g\in G$ and $\tilde f\colon\mathcal{Q}\to T_{[g]}(AdS_l)$ contains the main information while σ is the previously explained isometry. The action of $h\in\mathrm{SO}(1,l-1)$ on $\big([g],\tilde f\circ\sigma\big)$ is defined by means of any fixed isomorphism $\varphi_0\colon\mathrm{SO}(1,l-1)\to\mathrm{SO}(\mathcal{Q})$ by

$$\big([g],\tilde f\circ\sigma\big)\cdot h=\big([g],\tilde f\circ\varphi_0(h)\circ\sigma\big). \tag{5.229}$$

The map β is a morphism of principal bundle because

$$\beta[g,A]\cdot\varphi_0^{-1}(B)=\big([g],d\tau(g)A\circ\sigma\big)\cdot\varphi_0^{-1}(B)=\big([g],d\tau(g)A\circ B\circ\sigma\big)=\beta\big([g,A]\cdot B\big).$$

It remains to be proved that β is a bijection. Surjectivity is natural: since $d\tau(g)$ is an isometry, $d\tau(g)A$ runs over the whole SO $\big(T_{[g]}(AdS_l)\big)$ when A runs over $\mathrm{SO}(\mathcal{Q})$. Injectivity is as follows; let's suppose $\beta[g,A]=\beta[g',A']$. It is immediate that in this case, $\exists h\in H$ such that $g'=gh$. Using the fact that $d\pi\circ dR_{h^{-1}}\circ d\pi^{-1}=\mathrm{id}$ and $d\tau(h)d\pi=d\pi dL_h$, we have

$$d\tau(g)_{\vartheta}A=d\tau(g)d\tau(h)A'=d\tau(g)d\tau(h)d\pi dR_{h^{-1}}d\pi^{-1}A'=d\pi\,\mathrm{Ad}(h)d\pi^{-1}A'=\alpha(h)A'.$$

□

From now on, we identify $G\times_\alpha\mathrm{SO}(\mathcal{Q})$ with the frame bundle over AdS_l.

5.13.2 Spin structure

We consider the principal bundle

$$\begin{tikzcd} \mathrm{Spin}(1,l-1) \arrow[r, squiggly] & G\times_{\tilde\alpha}\mathrm{Spin}(1,l-1) \arrow[d, "p"] \\ & G/H \end{tikzcd} \tag{5.230}$$

where $\times_{\tilde\alpha}$ is the following equivalence relation on $G\times\mathrm{Spin}(1,l-1)$. We say that $(g,s)\sim(g',s')$ if and only if there exists a $h\in H$ such that

(i) $g'=gh$,

(ii) $\chi(s')=\mathrm{Ad}(h^{-1})\chi(s)$.

Notice that the second condition implies that $\mathrm{Ad}(h)\in\mathrm{SO}_0(\mathcal{Q})$. It is easy to prove that the given structure is well defined and is a principal bundle. Now we consider the spin structure as follows:

$$\begin{tikzcd} \mathrm{Spin}(1,l-1) \arrow[r, squiggly] & G\times_{\tilde\alpha}\mathrm{Spin}(1,l-1) \arrow[rr, "\varphi"] \arrow[dr] & & G\times_\alpha\mathrm{SO}(\mathcal{Q}) \arrow[dl] & \mathrm{SO}(\mathcal{Q}) \arrow[l, squiggly] \\ & & G/H & & \end{tikzcd} \tag{5.231}$$

where $\varphi[g,s]=[g,\chi(s)]$. It is well defined since when $[g,s]=[g',s']$, there exists a $h\in H$ with $\chi(s')=\mathrm{Ad}(h^{-1})\chi(s)$ such that $\varphi[g',s']=\varphi[gh,s']=[gh,\chi(s')]=[gh,\mathrm{Ad}(h^{-1})\chi(s)]=[g,\chi(s)]=\varphi[g,s]$.

5.13.3 Reduction of the structural group

The case of AdS_l can be seen in the setting of subsection **??**. Let us show now that the bundle

$$\begin{tikzcd} H_0 \arrow[r, squiggly] & G \arrow[d, "\pi"] \\ & G/H \end{tikzcd} \tag{5.232}$$

is a reduction to H_0 (the identity component of $\mathrm{SO}(\mathcal{Q})$) of

$$\begin{tikzcd} G \arrow[r, squiggly] & r(G) \arrow[d, "\pi"] \\ & G/H. \end{tikzcd} \tag{5.233}$$

Indeed, $u\colon G \to r(G)$ given by $u(g) = r(g)$ provides the reduction homomorphism: $r(gh)X = d\pi dL_{gh}X$ while $\big(r(g)\cdot h\big)X$ is the same.

Lemma 5.60.
The tangent space $T(G/H)$ is an associated bundle of $r(G)$ trough the identification

$$\begin{aligned} \beta'\colon r(G) \times_\rho \mathcal{Q} &\to T(G/H) \\ [r(g), X] &\mapsto r(g)X \end{aligned} \tag{5.234}$$

where $\rho(h)X = \mathrm{Ad}(h)X$, so that the quotient is given by $[g, X] = [gh, \mathrm{Ad}(h^{-1})X]$.

Proof. The proof is entirely similar to the one of lemma **??**. □

Chapter 6

Relativistic fields and group theory

6.1 Mathematical framework of field theory

This is a short review; the aim is to see why the quantum theory of fields needs representations of the Poincaré group. It will be mostly physics oriented. References dealing with field theory including gauge theory and representations are [31, 44, 44, 48, 51–55].

6.1.1 Axioms of the (quantum) relativistic field theory

The quantum mechanics is based on a few number of axioms:

(i) We have a Hilbert space $\mathscr{H}$. A physical state is given by a **ray** in $\mathscr{H}$, i.e. a set

$$\mathcal{R} = \{\xi\psi : |\xi| = 1\}$$

for a certain $\psi \in \mathscr{H}$ with $\langle\psi|\psi\rangle = 1$. In other words, the set of physical sates is the quotient of the set of unital vectors in $\mathscr{H}$ by the relation $\psi \sim \psi'$ if and only if $\psi = \xi\psi'$ for some unimodular complex number ξ. We denote by $\operatorname{Ray}\mathscr{H}$ the set of all rays in $\mathscr{H}$.

(ii) The observables are represented by hermitian linear operators on $\mathscr{H}$. A state $\mathcal{R}$ has value α for the observable A if $A\mathcal{R} = \alpha\mathcal{R}$, where the action of A on the ray is obvious (and well defined because A is linear).

(iii) If one has a system described by a state $\mathcal{R}$, and if one want to measure if it is in one of the state $\mathcal{R}_1, \ldots, \mathcal{R}_n$ (orthogonal rays), the answer will be $\mathcal{R}_i$ with probability

$$P(\mathcal{R} \to \mathcal{R}_i) = |\langle\mathcal{R}|\mathcal{R}_i\rangle|^2.$$

If the $\mathcal{R}_n$ form a complete system, one has a theorem which states that

$$\sum_i P(\mathcal{R} \to \mathcal{R}_i) = 1.$$

(iv) The rays of $\mathscr{H}$ furnish a representation of the (identity component of) Poincaré group.

This last point can look strange; we will see later (page 240) how it comes. It is the expression of a relativistic theory. That axiom is the reason why one make intensive use of representation theory in relativistic (quantum) field theory . . . or maybe the intensive use of representation theory is the reason of that axiom. However, we will make an intensive use of representation theory developed in chapter **??**.

6.1.2 Symmetries and Wigner's theorem

Consider the following situation: someone observes a system in a state $\mathcal{R}$, and makes measures $P(\mathcal{R} \to \mathcal{R}_i)$. An other person observes the same system which is, for him, in a state $\mathcal{R}'$ and observes $P(\mathcal{R}' \to \mathcal{R}'_i)$.

If two observers are related by a transformation of the Hilbert state which induces $\mathcal{R} \to \mathcal{R}'$ and $\mathcal{R}_i \to \mathcal{R}'_i$, there are said **equivalent** if

$$P(\mathcal{R} \to \mathcal{R}_i) = P(\mathcal{R}' \to \mathcal{R}'_i). \tag{6.1}$$

Let us say it more precisely from a mathematical point of view. A **symmetry** is an invertible operator $T\colon \operatorname{Ray}\mathscr{H} \to \operatorname{Ray}\mathscr{H}$ such that for any $\phi_i \in \mathcal{R}_i$, $\phi'_i \in T\mathcal{R}_i$ and $\phi''_i \in T^{-1}\mathcal{R}_i$,

$$|\langle\phi'_1|\phi'_2\rangle|^2 = |\langle\phi_1|\phi_2\rangle|^2 = |\langle\phi''_1|\phi''_2\rangle|^2 \tag{6.2}$$

Remark 6.1.
Here, neither $\mathcal{R}$ nor $\mathcal{R}'$ are measurable: the P's only are measurable.

The following can be found in [44] p.91, [51] p.354.

Theorem 6.2 (Wigner)**.**
Any symmetry T is induced by an operator U on $\mathscr{H}$ such that $\psi \in \mathcal{R}$ implies $U\psi \in \mathcal{R}'$. This operator is either unitary and linear, either anti-unitary and antilinear.

So, the symmetry operator must satisfy

$$\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle \tag{6.3a}$$
$$U(\xi\psi + \eta\psi) = \xi U\psi + \eta U\phi, \tag{6.3b}$$

or

$$\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle^* \tag{6.4a}$$
$$U(\xi\psi + \eta\psi) = \xi^* U\psi + \eta^* U\phi. \tag{6.4b}$$

In the anti-linear case operator, we do not define $U^\dagger$ by $\langle\phi|U^\dagger\psi\rangle = \langle U\phi|\psi\rangle$ because the left-hand side should be anti-linear with respect to ψ while the right-had should be linear. In place, for an antilinear operator A, we define $A^\dagger$ by

$$\langle\phi|A^\dagger\psi\rangle = \langle A\phi|\psi\rangle^* = \langle\psi|A\phi\rangle. \tag{6.5}$$

In this way, the definitions of unitary and anti-unitary in term of dagger are the same: $U^\dagger = U^{-1}$.

For any transformation $T\colon \operatorname{Ray}\mathscr{H} \to \operatorname{Ray}\mathscr{H}$, the Wigner's theorem provides an operator $U(T)\colon \mathscr{H} \to \mathscr{H}$ which induces T on Ray. If the operator T depends on a parameter θ, the operator $U(T(\theta))$ depends on θ. If T depends continuously on the parameter then the family $U(T(\theta))$ only contains unitary/linear operators or only antiunitary/antilinear operators.

In physical cases, $T(\theta)$ is mostly a Poincaré transformation: $\theta = (\Lambda, p)$. But $T(\mathbb{1}, 0)$ is the identity which is represented by $U(\mathbb{1}, 0) = \mathbb{1}$. Then all the (connected to identity) Poincaré transformations are represented by linear and unitary operators on $\mathscr{H}$.

We will follow the proof given in [44]. An other form of the proof can be found in [51]. The latter use a slightly different formalism in the axioms of the quantum mechanics; this is explained in appendix A.1. It is now time to prove the theorem.

Proof of Wigner's theorem. We consider an orthonormal basis $\{\psi_k\}$ of $\mathscr{H}$ with $\psi_k \in \mathcal{R}_k$, and a choice of $\psi'_k \in T\mathcal{R}_k$. From this and the assumptions, we have

$$|\langle\psi'_k|\psi'_l\rangle|^2 = |\langle\psi_k|\psi_l\rangle|^2 = \delta_{kl}.$$

Then $\langle\psi'_k|\psi'_k\rangle = 0$ whenever $k \neq l$ and, since $\langle\psi'_k|\psi'_k\rangle$ is real and positive, $\langle\psi'_k|\psi'_k\rangle = 1$. So $\langle\psi'_k|\psi'_l\rangle = \delta_{kl}$.

The set ψ'_k is also complete in $\mathscr{H}$. Indeed suppose that we have a vector $\psi' \in \mathscr{H}$ such that $\langle\psi'|\psi'_k\rangle = 0$ for all k. If $\psi' \in \mathcal{R}$, we consider a $\psi'' \in T^{-1}\mathcal{R}$ and we have

$$|\langle\psi''|\psi_k\rangle|^2 = |\langle\psi'|\psi'_k\rangle|^2 = 0,$$

which contradicts the fact that the ψ_k's form a complete set. Now we have to fix a phase convention for the ψ_k. Since there are no canonical choice of phase, we fix with respect to an arbitrary one of the ψ_k, say ψ_1. We put

$$\gamma_k = \frac{1}{\sqrt{2}}(\psi_1 + \psi_k) \in \mathcal{C}_k \tag{6.6}$$

for $k \neq 1$. Any $\gamma'_k \in T\mathcal{C}_k$ can be written in the basis $\{\psi'_k\}$:

$$\gamma'_k = \sum_l c_{kl}\psi'_l. \tag{6.7}$$

From assumption (6.1) and the fact that $|c_{kl}|^2 = |\langle\gamma'_k|\psi'_l\rangle|^2$, we find, for $k, l \neq 1$

$$|c_{kl}|^2 = \frac{1}{2}\delta_{kl}.$$

We can choose the phase of γ'_k and ψ'_k in order to get $c_{kk} = c_{k1} = 1/\sqrt{2}$. For this, we begin to fix γ'_k in such a manner to get $c_{k1} = 1/\sqrt{2}$ (from $|c_{k1}| = |\langle\gamma'_k|\psi'_1\rangle|$), and next we fix ψ'_k for the c_{kk}. From now on, the so chosen γ'_k and ψ'_k are denoted by $U\gamma_k$ and $U\psi_k$.

What we did until now is to take a basis $\{\psi_k\}$ of $\mathscr{H}$ and define $\gamma_k = 1/\sqrt{2}(\psi_1 + \psi_k)$. Next we had chosen the phases of $\psi'_k \in T\mathcal{R}_k$ and $\gamma'_k \in T\mathcal{C}_k$ in order to have

$$\begin{aligned} c_{kk} = c_{k1} &= 1/\sqrt{2} \quad \forall k, \\ c_{kl} &= 0 \qquad \text{if } l \neq k \text{ and } l \neq 1. \end{aligned} \tag{6.8}$$

This allows us to check a certain linearity for the operator U:

$$\begin{aligned} U\left(\frac{1}{\sqrt{2}}(\psi_k + \psi 1)\right) &= U\gamma_k \\ &= \gamma'_k \\ &= \frac{1}{\sqrt{2}}\psi'_1 + \frac{1}{\sqrt{2}}\psi'_k \qquad \text{from (6.7) and (6.8)} \\ &= \frac{1}{\sqrt{2}}\big(U\psi_1 + U\psi_k\big). \end{aligned} \tag{6.9}$$

Now we have to build U on a general vector $\psi = \sum_k \psi_k \in \mathcal{R}$. Any vector $\psi' \in T\mathcal{R}$ can be decomposed with respect to the basis $\{\psi'_k = U\psi_k\}$:

$$\psi' = \sum_k C'_k U\psi_k. \tag{6.10}$$

From the conservation of probability $|\langle\psi_k|\psi\rangle|^2 = |\langle U\psi_k|\psi'\rangle|^2$ and $|\langle\gamma_k|\psi\rangle|^2 = |\langle U\gamma_k|\psi'\rangle|^2$, we find

$$|C_k|^2 = |C'_k|^2, \tag{6.11a}$$

$$|C_k + C_1|^2 = |C'_k + C'_1|^2. \tag{6.11b}$$

If one writes $C_k = a_k + ib_k$, one finds $\mathrm{Re}(C_k/C_1) = (a_k a_1 + b_k b_1)/|C_1|^2$. By doing the same with C'_k and using (6.11),

$$\mathrm{Re}(C_k/C_1) = \mathrm{Re}(C'_k/C'_1). \tag{6.12}$$

Equation (6.11a) also imposes

$$|C_k/C_1|^2 = |C'_k/C'_1|^2, \tag{6.13}$$

while compatibility between (6.13) and (6.12) requires

$$\mathrm{Im}(C_k/C_1) = \pm\,\mathrm{Im}(C'_k/C'_1). \tag{6.14}$$

Equations (6.12) and (6.14) show that C_k and C'_k must satisfy

$$C_k/C_1 = C'_k/C'_1 \tag{6.15a}$$

xor

$$C_k/C_1 = (C'_k/C'_1)^*. \tag{6.15b}$$

For a given ψ we have to show that the choice must be the same for all the C_k[1]. Let $l \neq k$ and suppose that $C_k/C_1 = C'_k/C'_1$ and $C_l/C_1 = (C'_l/C'_1)^*$; we will show that in this case, one of the two ratios is real. So we can suppose $k \neq 1 \neq l$. We consider the vector $\Phi = \frac{1}{\sqrt{3}}(\psi_1 + \psi_k + \psi_l)$,

$$\Phi' = \frac{\alpha}{\sqrt{3}}\big(U\psi_1 + U\psi_k + U\psi_l\big)$$

where $\alpha \in \mathbb{C}$ satisfies $|\alpha| = 1$. The conservation of probability $|\langle\Phi|\psi\rangle|^2 = |\langle\Phi'|\psi'\rangle|^2$ gives $|C_1 + C_k + C_l|^2 = |C'_1 + C'_k + C'_l|^2$. Since $|C_1|^2 = |C'_1|^2$, we can divide the left hand side by $|C_1|^2$ and the right one by $|C'_1|^2$. We find

$$\left|1 + \frac{C_k}{C_1} + \frac{C_l}{C_1}\right|^2 = \left|1 + \frac{C'_k}{C'_1} + \frac{C'_l}{C'_1}\right|^2.$$

Using the assumption $C_k/C_1 = C'_k/C'_1$ and $C_l/C_1 = (C'_l/C'_1)^*$, we are in a case of an equation of the form $|u+v|^2 = |u+v^*|^2$ with $u, v \in \mathbb{C}$. If we write $u = a + bi$ and $v = x + iy$, we find $b + y = \pm(b - y)$, so that it leaves the choice $y = 0$ or $b = 0$ which corresponds to $(C_k/C_1) \in \mathbb{R}$ or $(C_l/C_1) \in \mathbb{R}$. So the coefficients C'_k $(k \neq 1)$ in the expansion (6.10) must satisfy

$$C_k/C_1 = C'_k/C'_1 \quad \forall k \tag{6.16a}$$

[1] We will show later that for a given T, the choice must be the same for all the ψ.

xor

$$C_k/C_1 = (C'_k/C'_1)^* \quad \forall k. \tag{6.16b}$$

Note that the phase of C_1 is not yet fixed. We naturally choice $C_1 = C'_1$ or $C_1 = {C'_1}^*$ following the case. We define $U\colon \mathscr{H} \to \mathscr{H}$ by

$$U\Big(\sum_k C_k\psi_k\Big) = \sum_k C_k U\psi_k \qquad \text{if (6.16a)}, \tag{6.17a}$$

xor

$$U\Big(\sum_k C_k\psi_k\Big) = \sum_k C_k^* U\psi_k \qquad \text{if (6.16b)}. \tag{6.17b}$$

One can explicitly check that it preserves the probability because $|\langle\psi|\psi_k\rangle|^2 = |C_k|^2$ while $|\langle U\psi|U\psi_k\rangle|$ is equal to $|C_k|^2$ or $|C_k^*|^2$ (which are the same) following the case (6.17a) or (6.17b).

Now we have to prove that the choice (6.17a) or (6.17b) is fixed by the data of T and must be the same for all the $\psi \in \mathscr{H}$. For, let us consider two vectors $\phi = \sum A_k\psi_k$, $\varphi = \sum B_k\psi_k$ and suppose that

$$U\phi = \sum_k A_k U\psi_k \text{ but } U\varphi = \sum_k B_k^* U\psi_k.$$

In order to see that it is impossible, looks at the conservation of probability $|\sum_k A_k B_k^*|^2 = |\sum_k A_k B_k|^2$, then

$$\sum_{kl}\big(B_l^* B_k A_l A_k^* - B_l^* B_k A_l^* A_k\big) = \sum_{kl} B_l^* B_k \operatorname{Im}(A_l A_k^*) = 0. \tag{6.18}$$

Since $A_l A^* l \in \mathbb{R}$, we can regroup each term (k, l) with the corresponding term (l, k). We get

$$0 = \sum_{kl} \operatorname{Im}(A_l A_k^*)(B_l^* B_k - B_k^* B_l) = \sum_{kl} \operatorname{Im}(A_k^* A_l) \operatorname{Im}(B_k^* B_l). \tag{6.19}$$

We can find a vector $\sum_k C_k\psi_k$ such that

$$\sum_{kl} \operatorname{Im}(C_k^* C_l) \operatorname{Im}(A_k^* A_l) \neq 0 \tag{6.20a}$$

and

$$\sum_{kl} \operatorname{Im}(C_k^* C_l) \operatorname{Im}(B_k^* B_l) \neq 0. \tag{6.20b}$$

In order to see how to find such a vector, let us show that there always exists a choice (i, j) such that $B_i^* B_j$ is not real. Let us say $B_1 = x + iy$ and $B_k = a_k + bi$. If $y \neq 0$, the condition $\operatorname{Im}(B_1^* B_k) = 0$ gives $B_k = \frac{b_k}{y} B_1$. It is always possible to find a sequence (b_k) which gives 1 as norm for $\sum B_k\psi_k$; the problem is not there. The problem is that $B_k/B_1 \in \mathbb{R}$, so that the choice (6.17) is not a true choice. For the same reason, all the $B_i^* B_k$ can't be pure imaginary.

Now we can find the vector which satisfy (6.20). There are several cases. If there is a pair (k, l) such that $A_k^* A_l$ and $B_k^* B_l$ are both complex, we can take all C_i's zero for $k \neq i \neq l$ and choose C_k and C_l in such a way that $C_k^* C_l$ is not real. If there is a pair (k, l) with $A_k^* A_l$ complex and $B_k^* B_l$ real, we consider a pair (m, n) such that $B_m^* B_n$ is complex. If $A_m^* A_n$ is complex, we take all the C_i's zero except C_m and C_n such that $\operatorname{Im}(C_m^* C_n) \neq 0$. If $A_m^* A_n$ is real, we take all the C_i's zero except C_k, C_l, C_m, C_n which we choose in such a way that $\operatorname{Im}(C_m^* C_n) \neq 0$ and $\operatorname{Im}(C_k^* C_k) \neq 0$.

Equation (6.20a) makes that the same choice must be made for $\sum A_k\psi_k$ and $\sum C_k\psi_k$ (if it was not the case, we would have an equation of the form of (6.19)). For the same reason, the same choice must be made for $\sum B_k\psi_k$ and $\sum C_k\psi_k$. So we conclude that the data of T fixes the choice between (6.17a) and (6.17b) and that this choice must be the same for all the vectors of $\mathscr{H}$.

We have to show that the possibility (6.17a) makes U linear and unitary while the possibility (6.17b) makes U antilinear and antiunitary. For we consider $\psi = \sum_k A_k\psi_k$ and $\phi = \sum_k B_k\psi_k$. If (6.17a) works,

$$\begin{aligned}
U(\alpha\psi + \beta\phi) &= U\Big(\sum_k (\alpha A_k + \beta B_k)\psi_k\Big) \\
&= \sum_k (\alpha A_k + \beta B_k) U\psi_k \\
&= \alpha U\psi + \beta U\phi,
\end{aligned} \tag{6.21}$$

and

$$\langle U\psi|U\phi\rangle = \sum_{kl} A_k^* B_l \langle U\psi_k|U\psi_l\rangle = \sum_k A_k^* B_k, \tag{6.22}$$

so that $\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle$. Thus in this case U is linear and unitary. In the case where (6.17a) works, the computations are almost the same:

$$\begin{aligned} U(\alpha\psi + \beta\phi) &= U\Big(\sum_k(\alpha A_k + \beta B_k)\psi_k\Big) \\ &= \sum_k(\alpha^* A_k^* + \beta^* B_k^*)U\psi_k \\ &= \alpha^* U\psi + \beta^* U\phi, \end{aligned} \tag{6.23}$$

and

$$\langle U\psi|U\phi\rangle = \sum_{kl} A_k B_l^* \langle U\psi_k|U\psi_l\rangle = \sum_k A_k B_k^*, \tag{6.24}$$

so that $\langle U\psi|U\phi\rangle = \langle\psi|\phi\rangle^*$. In this case, U is antilinear and antiunitary.

□

6.1.3 Projective representations

If $T_1(\mathcal{R}_n) = \mathcal{R}_n'$ and $\psi_n \in \mathcal{R}_n$, then $U(T_1)\psi_n \in \mathcal{R}_n'$. If $T_2(\mathcal{R}') = \mathcal{R}''$, then $U(T_2)U(T_1)\psi_n \in \mathcal{R}_n''$. But $U(T_2T_1)\psi_n$ also belongs to $\mathcal{R}_n''$. Then there exists a $\phi_n(T_2,T_1) \in \mathbb{R}$ such that

$$U(T_2)U(T_1)\psi_n = e^{i\phi_n(T_2,T_1)}U(T_2T_1)\psi_n.$$

Note that for any $\psi \in \mathscr{H}$, there exists a $\lambda \in \mathbb{R}$ such that $\|\lambda\psi\| = 1$. Since a real can be sent out the $U(T)$'s, for *any* $\psi \in \mathscr{H}$, there exists a ϕ which only depends on $\psi/\|\psi\|$ such that

$$U(T_2)U(T_1)\psi = e^{i\phi(T_2,T_1)}U(T_2T_1)\psi \tag{6.25}$$

Proposition 6.3.
The ϕ doesn't depend at all on the ψ:

$$U(T_2)U(T_1) = e^{i\phi(T_2,T_1)}U(T_2T_1). \tag{6.26}$$

Proof. Let us consider a ψ_A and a ψ_B which are not proportional each other. One has a $\phi_{AB}(T_2,T_1)$ such that

$$\begin{aligned} e^{i\phi_{AB}(T_2,T_1)}U(T_2T_1)(\psi_A+\psi_B) &= U(T_2)U(T_1)(\psi_A+\psi_B) \\ &= e^{i\phi_A(T_2,T_1)}U(T_2T_1)\psi_A \\ &\quad + e^{i\phi_B(T_2,T_1)}U(T_2T_1)\psi_B. \end{aligned} \tag{6.27}$$

Now, we apply $U(T_2T_1)^{-1}$ to both sides. If it is unitary, the $e^{i\phi}$ get out without problems; else is get out as $e^{-i\phi}$:

$$e^{\pm i\phi_{AB}}(\psi_A+\psi_B) = e^{\pm i\phi_A}\psi_A + e^{\pm i\phi_B}\psi_B. \tag{6.28}$$

Since ψ_A and ψ_B are linearly independent, the only solution is $e^{i\phi_{AB}} = e^{i\phi_A} = e^{i\phi_B}$.

□

Since the operators $U(T)$ must only fulfil

$$U(T_2)U(T_1) = e^{i\phi(T_2,T_1)}U(T_2,T_1), \tag{6.29}$$

these form a **projective representation** of the symmetry group on the physical Hilbert space $\mathscr{H}$.

Remark 6.4.
In order to have some physical relevance, this demonstration supposes that a state $\psi_A+\psi_B$ exists in nature. If one can divide the particles in several "incompatibles" classes labeled by a,b such that $\psi_a+\psi_b$ doesn't exist, then equation (6.29) is false and one has to write

$$U(T_2)U(T_1)\psi_a = e^{i\phi_a(T_2,T_1)}U(T_2T_1)\psi_a$$

because we can't show that $\phi_a = \phi_b$ from the simple fact that $\psi_a+\psi_b$ doesn't exist !

For example, physicists think that there are no superposition of state of integer and semi-integer spin.

Remark 6.5.
If the group satisfies some requirements, one can choose $\phi = 0$. From now we suppose that we are in this case: we work with "true" representations.

6.1.4 Representations and power expansions

Let G be an arc connected Lie group whose elements are denoted by $T(\theta)$ with θ, a continuous family of parameters (from a local chart). The multiplication law is given by a function $f\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$:

$$T(\theta')T(\theta) = T\big(f(\theta',\theta)\big). \tag{6.30}$$

If $\theta = 0$ is the coordinate of the identity,

$$f(0,\theta) = f(\theta,0) = \theta. \tag{6.31}$$

We suppose that G acts on the rays of a Hilbert space $\mathscr{H}$, so that there are represented on $\mathscr{H}$ by unitary operators $U\big(T(\theta)\big)$. We denote by W the group of transformations of $\mathscr{H}$; roughly speaking,

$$W = U(G).$$

Now, we are going to cheat a little. We know that there exists a normal neighbourhood of e in W. In simple words, the map $\exp\colon \mathcal{W} \to W$ is a diffeomorphism between the elements of $\mathcal{W}$ "close" to 0 and the ones of W close to e. By *close to*, we mean that the components of θ are small enough. If $\{it_a\}$ is a basis of $\mathcal{W}$, we define

$$U(T(\theta)) = e^{i\theta^a t_a}. \tag{6.32}$$

In other words, one considers the exponential map for a neighbourhood of identity.

The cheat is the fact that $U(T(\theta))$ is actually defined by Wigner's theorem from the data of the group G. So equation (6.32) should be seen as a requirement in the choice of the basis $\{t_a\}$.

Remark 6.6.
The i in the exponential in (6.32) and in the definition of the basis $\{it_a\}$ is a convention in order the t_a's to be hermitian. Indeed, the Lie algebra of a group of unitary matrices is made of anti*hermitian matrices.*

With all that,

$$U(T(\theta)) = \mathbb{1} + i\theta^a t_a + \frac{1}{2}\theta^b\theta^c t_{bc} + \ldots \tag{6.33}$$

where t_{bc} is defined (among other requirements) to absorb the "intuitive" minus sign in the third term.

Now we are going to explore some consequences of equation (6.30). Equation (6.31) makes the expansion of f as

$$f^a(\theta',\theta) = \theta^a + {\theta'}^a + f^a_{bc}{\theta'}^b\theta^c + \ldots \tag{6.34}$$

From expansions (6.34) and (6.33) of f and $U(T(\theta))$, "group structure" equation (6.30) gives (at order two):

$$t_{bc} = -t_b t_c - i f^a_{bc} t_a \tag{6.35}$$

and nothing for the first order. Then, providing that one knows the group structure (the f), one knows the second order of the representation from the first one. From equation (6.32), one finds the value of t_{ab}:

$$e^{i\theta^a t_a} = 1 + i\theta^a t_a + \frac{1}{2}(i)^2(\theta^a t_a)(\theta^b t_b),$$

up to constant coefficients, one can choose t_{ab} to be symmetric with respect to a and b:

$$t_{ab} = \frac{1}{2}(t_a t_b + t_b t_a).$$

Taking this convention and computing $t_{bc} - t_{cb}$ from (6.35), we find

$$[t_a, t_b] = iC^c_{ab} t_c \tag{6.36}$$

with $C^c_{ab} = f^c_{ab} - f^c_{ba}$.

On the other hand, one knows that if a group is abelian, its algebra is also abelian; we can see it here by considering that if G is abelian, $f(\theta,\theta') = f(\theta',\theta)$, then f^c_{ab} is symmetric and $[t_a, t_b] = 0$. We can say more about f Since the t_a commute, equations (6.30) and (6.32) make that

$$e^{if(\theta,\theta')^a t_a} = e^{i\theta^a t_a} e^{i{\theta'}^b t_b} = e^{i(\theta^a + {\theta'}^a)t_a}, \tag{6.37}$$

so that

$$f(\theta,\theta') = \theta + \theta'.$$

6.2 The symmetry group of nature

6.2.1 Spin and double covering

Some of literature carry an ambiguity in the choice of the right space-time symmetry group in the quantum field theory. A very good and deep discussion about the choice of the space-time symmetry group of nature is given in the book [31] which will be used here. An other enlightening review can be found in [56].

From a relativistic point of view, the group is the Poincaré group of all the maps $\mathbb{R}^4 \to \mathbb{R}^4$ which leaves invariant the quantity $s^2 = -t^2 + x^2 + y^2 + z^2$. At this point we can already make an important remark: the so defined quantity s is in fact *not* a relativistic invariant. Indeed if I follow a (spatially) closed path, I will measure $\Delta t \neq 0$ and $\Delta x = \Delta y = \Delta z = 0$ because in *my* frame, my displacement is zero. A guy who keeps at my starting point will measure (between the beginning and the end of my travel) $\Delta' t \neq 0$ and also $\Delta x = \Delta y = \Delta z = 0$. If $s = s'$, then $\Delta t = \Delta t'$.

So the relativistic invariance is only local: $ds^2 = ds'^2$, and as far as relativity is concerned, one can work with infinitesimal transformations only. In this case, the distinction between the *groups* $L^{\uparrow}_{+}$ and $\mathrm{SL}(2,\mathbb{C})$ is no relevant. Intuitively, we choose $L^{\uparrow}_{+}$ to be the space-time symmetry group. As we will see the difference will reveal to be crucial in relativistic field theory because $L^{\uparrow}_{+}$ has no half-integer spin representations.

This group naturally splits into two parts: the translations and the rotations (and boost). As far as I know, the translation part makes no difficulties. For the other one, there are some difficulties to find the *minimal* group of symmetry. First, one often want to separate the space-time inversions P and T from the remaining: the group then becomes the homogeneous orthochrone Lorentz group $L^{\uparrow}_{+}$. An other often presented group is $\ldots\mathrm{SL}(2,\mathbb{C})$. This is our choice here. The physical reason of this choice is all but immediate. As we will see during the following pages, an elementary particle is an irreducible representation of the symmetry group.

For massive particles, the relevant subgroup of $\mathrm{SL}(2,\mathbb{C})$ reveals to be $SU(2)$. If we had chosen the most intuitive $L^{\uparrow}_{+}$, we would have found $\mathrm{SO}(3)$. There is an important difference between $SU(2)$ and $\mathrm{SO}(3) = SU(2)/\mathbb{Z}_2$: the first one admits representations of any integer and half-integer spin while the second only posses the integer spin representations (cf. page **??**).

Let us now be more precise about the relation between $L^{\uparrow}_{+}$ and $\mathrm{SL}(2,\mathbb{C})$. A know result is

$$L^{\uparrow}_{+} = \frac{\mathrm{SL}(2,\mathbb{C})}{\mathbb{Z}_2}.$$

Let $\mathrm{Spin}\colon \mathrm{SL}(2,\mathbb{C}) \to L^{\uparrow}_{+}$ be the surjective homomorphism with kernel $\pm\mathbb{1}_{2\times 2}$ giving this relation. We will not give a complete proof, but we will explain how $\mathrm{SL}(2,\mathbb{C})$ acts by isometries on $\mathbb{R}^4$. First, we remark that there exists a bijection between $\mathbb{R}^4$ and the 2×2 complex hermitian matrices:

$$v = \begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix} = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}. \tag{6.38}$$

If $\lambda \in \mathrm{SL}(2,\mathbb{C})$, the matrix $\lambda v \lambda^{\dagger}$ is also hermitian and $\|v\|^2 = \det v$. Thus

$$\begin{aligned} \Lambda(\lambda)\colon \mathbb{R}^4 &\to \mathbb{R}^4 \\ v &\mapsto \lambda v \lambda^{\dagger} \end{aligned} \tag{6.39}$$

is a Lorentz transformation if and only if $|\det\lambda| = 1$. Moreover,

$$\Lambda(\lambda\lambda') = \Lambda(\lambda)\Lambda(\lambda').$$

If $\lambda' = e^{\phi}\lambda$, then $\Lambda(\lambda') = \Lambda(\lambda)$, thus it is natural to impose $\det v = 1$ and to consider $\mathrm{SL}(2,\mathbb{C})$ instead of $L(2,\mathbb{C})$ to fit $L^{\uparrow}_{+}$. Now, $\Lambda(\lambda) = \Lambda(-\lambda)$, and we wish to consider $\mathrm{SL}(2,\mathbb{C})/\mathbb{Z}_2$.

I think the problem is the following: as far as the action of the "nature group" on the space-time is concerned, it is sufficient to consider $L^{\uparrow}_{+}$. But the group which acts on the state space is wider: it must be $SL(2,\mathbb{C})$.

From now, when we say "Poincaré group", we mean $\mathrm{SL}(2,\mathbb{C}) \times \mathbb{R}^4$ while "Lorentz" means $\mathrm{SL}(2,\mathbb{C})$ acting on $\mathbb{R}^4$ by $\Lambda(\lambda)v = \lambda v \lambda^{\dagger}$.

Let us continue the discussion of page **??**. A know result is the fact that the map Spin restricts to a surjective homomorphism $\mathrm{Spin}\colon SU(2) \to \mathrm{SO}(3)$ with kernel $\pm\mathbb{1}$ giving the relation $\mathrm{SO}(3) = SU(2)/\mathbb{Z}_2$. If one considers a representation $\rho\colon \mathrm{SO}(3) \to GL(V)$, then $\tilde{\rho} = \rho \circ \mathrm{Spin}$ is a representation of $SU(2)$ on V. So every representation of $\mathrm{SO}(3)$ comes from a representation of $SU(2)$.

As far as the transformation rule of a (quantum mechanical) wave function under a rotation $R \in \mathrm{SO}(3)$ is concerned, one can see (it is done in [31]) that the try

$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \to T(R) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

doesn't works if $T(R)$ is a representation of SO(3) on $\mathbb{C}^2$. If one allows T to be a representation of $SU(2)$, then our choice —for an electron— should naturally be the spin one half representation $T = D^{(1/2)}$. Let us do it. The remaining problem is the following. Let's consider that in a certain frame, an electron is described by the wave function $\begin{pmatrix} \psi_1 & \psi_2 \end{pmatrix}$, the question is to know the wave function observed by a guy which use another frame linked to the first frame by $R \in \mathrm{SO}(3)$. We always have exactly two elements in SU(2) projected to R by Spin; namely $\mathrm{Spin}(\pm g) = R$; so how to choose between

$$D^{(1/2)}(g) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \quad \text{and} \quad D^{(1/2)}(-g) \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \; ?$$

The trick is to remark that a change of frame is not the mathematical process described by a single element R of SO(3), but a physical *continuous* process which begins at the identity and stops at R. In other word, we have to ask ourself *how to go from a frame to another* ? Taking as example the rotations around the x axis, we can look at two different path in SO(3) from $\mathbb{1}$ to $\mathbb{1}$ given by the same expression

$$R_1(t) = R_2(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix},$$

but considering $t\colon 0 \to 2\pi$ for R_1 and $t\colon 0 \to 4\pi$ for R_2. The covering map $\mathrm{Spin}\colon SU(2) \to \mathrm{SO}(3)$ allows us to lift any path in SO(3) to a path in $SU(2)$ in an unique way providing a starting point. In other words, if $\mathrm{Spin}(g) = R$,

$$\exists!\, \tilde{R}(t) \in SU(2) \text{ such that } \mathrm{Spin} \circ \tilde{R} = R \text{ and } \tilde{R}(0) = \mathbb{1},$$
$$\exists!\, \tilde{R}(t) \in SU(2) \text{ such that } \mathrm{Spin} \circ \tilde{R} = R \text{ and } \tilde{R}(0) = -\mathbb{1}.$$

The question is now: how to choose the right path among these two ? The answer comes from the homotopy of SO(3): the path R_1 and R_2 belongs to two different classes.

Considering the "change of frame" as a continuous process, the initial point is naturally chosen to be $\mathbb{1}$. With this choice, the lift of R_1 and R_2 are given by

$$g_1(t) = g_2(t) = \begin{pmatrix} \cos\frac{t}{2} & -i\sin\frac{t}{2} \\ -i\sin\frac{t}{2} & \cos\frac{t}{2} \end{pmatrix}$$

with $t\colon 0 \to 2\pi$ for g_1 and $t\colon 0 \to 4\pi$ for g_2. In $SU(2)$, the ending point of g_1 is $-\mathbb{1}$ while the one of g_2 is $\mathbb{1}$.

It is still possible to say a lot of interesting thinks about the space-time symmetry group of nature; let's just conclude saying that $SU(2)$ is more adapted to the rotations of non zero spin than SO(3). (it is not intuitive !)

6.2.2 How to implement the Poincaré group

We are not making physics here, but differential geometry and group theory; so we will not discuss the physical relevance of the Poincaré group from a "speed of light" point of view. We consider the **Poincaré group** as the group of all the affine isometries of metric $\eta = diag(-1,1,1,1)$ and the **Lorentz group** as the subgroup of rotations and boost.

A Poincaré transformation of $\mathbb{R}^4$ is given by (Λ, a) with Λ a 4×4 matrix and $a \in \mathbb{R}^4$, a translation vector. The composition of (Λ, a) with (Λ', a') is given by $(\Lambda'\Lambda, \Lambda' a + a')$, the inverse is $(\Lambda^{-1}, -\Lambda^{-1}a)$, the neutral is $(\mathbb{1}, 0)$, and $(\det \Lambda)^2 = 1$.

The axiom (iv) at page 233 gives us a group of transformation of the rays in $\mathscr{H}$ parametrised by (Λ, a) such that

$$T(\Lambda', a')T(\Lambda, a) = T(\Lambda'\Lambda, \Lambda' a + a'), \tag{6.40}$$

$T(\Lambda, a)\colon \mathrm{Ray}\,\mathscr{H} \to \mathrm{Ray}\,\mathscr{H}$. Then Wigner's theorem defines a representation of the Poincaré group on $\mathscr{H}$ by unitary matrices :

$$\psi \to U(\Lambda, a)\psi.$$

Remark 6.7.
Wigner only ensure existence of projective *representations. Here we suppose that our symmetry group (maybe*

slightly different that Poincaré) is such that any projective representations can be turn into a classical representation. We will therefore use the composition law

$$U(\Lambda', a')U(\Lambda) = U(\Lambda'\Lambda, \Lambda' a + a') \tag{6.41}$$

instead of $U(\Lambda', a')U(\Lambda, a') = e^{i\phi(\Lambda,a,\Lambda',a')}U(\Lambda'\Lambda, \Lambda' a + a')$.

By axiom, the (connected) Poincaré group acts on rays of $\mathscr{H}$, and we have the representation U which form a group acting on $\mathscr{H}$. The Lie algebra acts also :

$$u\psi = \frac{d}{dt}\Big[U(t)\Big]_{t=0}\psi := \frac{d}{dt}\Big[U(t)\psi\Big]_{t=0}. \tag{6.42}$$

This definition is natural because $\mathscr{H}$ is a vector space: it can be identified with its tangent space: $U(t)\psi$ is a path in $\mathscr{H}$ and its derivative at $t = 0$ is still a well defined element in $\mathscr{H}$. Now recall that the operators U are unitary, so that the corresponding operators u are hermitian (therefore diagonalisable).

Let us consider an abelian subgroup A of Poincaré with Lie algebra $\mathfrak{a}$. One can find a basis of $\mathscr{H}$ made of common eigenvectors of a basis of $\mathfrak{a}$. In other words, one can find a basis of $\mathscr{H}$ which simultaneously diagonalises all $\mathfrak{a}$. If $\{a_i\}$ is a basis of $\mathfrak{a}$, one can find a basis $\{|\psi_\lambda\rangle\}$ (here λ labels a basis of $\mathscr{H}$: it might take continuous values) such that

$$a_i|\psi_\lambda\rangle = \lambda_i|\psi_\lambda\rangle. \tag{6.43}$$

6.2.3 Momentum operator

Of course, there exists an abelian subgroup of Poincaré: the pure translations, $A = \{U(\mathbb{1}, a)\}$. A basis of the Lie algebra is given by four vectors labeled as P^μ and defined by

$$P^\mu = \frac{d}{dt}\Big[U(\mathbb{1}, te^\mu)\Big]_{t=0}$$

where e^μ is the unit vector following the direction μ (for $\mu = 0$, $e^0 = (1, 0, 0, 0)$). One can consider a basis which diagonalises the P^μ's:

$$P^\mu|p, \sigma\rangle = p^\mu|p, \sigma\rangle \tag{6.44}$$

where by definition,

$$P^\mu|p, \sigma\rangle = \frac{d}{dt}\Big[U(\mathbb{1}, te^\mu)|p, \sigma\rangle\Big]_{t=0}. \tag{6.45}$$

Remark 6.8.
Be careful on a point: we don't say anything about the symbol "p" in the ket. The only property is that it labels a Hilbert space $\mathscr{H}$. But nothing is already imposed to $\mathscr{H}$: it must just carry a representation of the Poincaré group on its rays. In particular, it is a priori *false to say that p is a "momentum 4-vector" and that p^μ is a component of p. Naturally, our notations are adapted to think that ! Maybe it is a pedagogical mistake; I don't know.*

This remark can be disturbing: why is generally $|p, \sigma\rangle$ called "a state of momentum p" ? Since $U(\mathbb{1}, a)$ is unitary, P^μ is hermitian; the p^μ are eigenvalues for an hermitian operator, so by axiom (ii) (page 233) they are candidate to be physical values. But equation (6.45) shows that P^μ is what a physicist should call an "infinitesimal translation", so that Noether suggests us to interpret the eigenvalue as momentum. We are safe !

The parameters σ are not yet defined neither. It will come later. For the moment, we include into the definition of a **one particle state** that σ takes discrete values.

Since $U(\mathbb{1}, a) = e^{a_\mu P^\mu}$,

$$U(\mathbb{1}, a)|p, \sigma\rangle = e^{ia_\mu p^\mu}|p, \sigma\rangle.$$

Now we are interested in the determination of $U(\Lambda, a)|p, \sigma\rangle$.

Proposition 6.9.
The operators P^μ are subject to the "transformation law"

$$U(\Lambda, a)P^\mu U(\Lambda, a)^{-1} = \Lambda^\mu_\nu P^\nu. \tag{6.46}$$

Proof. Since operators $U(\Lambda, a)$ are linear, they can be putted in the derivative which defines P^μ. Using the composition law (6.41) we find :

$$\begin{aligned} U(\Lambda, a)P^\mu U(\Lambda, a)^{-1} &= \frac{d}{dt}\Big[U(\Lambda, a)U(\mathbb{1}, te^\mu)U(\Lambda, a)^{-1}\Big]_{t=0} \\ &= \frac{d}{dt}\Big[U(\mathbb{1}, t\Lambda e^\mu)\Big]_{t=0}. \end{aligned} \tag{6.47}$$

The Λ can be putted out of derivative; let us see it for a sum of two terms (here it is four) :

$$\begin{aligned}\frac{d}{dt}\Big[U(\mathbb{1},t(e^\mu+e^\nu))\Big]_{t=0}&=\frac{d}{dt}\Big[U(\mathbb{1},te^\mu)U(\mathbb{1},te^\nu)\Big]_{t=0}\\&=\frac{d}{dt}\Big[U(\mathbb{1},te^\mu)U(\mathbb{1},0)\Big]_{t=0}+\frac{d}{dt}\Big[U(\mathbb{1},0)U(\mathbb{1},te^\nu)\Big]_{t=0}\\&=P^\mu+P^\nu.\end{aligned}\tag{6.48}$$

Thus

$$\frac{d}{dt}\Big[U(\mathbb{1},\Lambda^\mu_\nu e^\nu)\Big]_{t=0}=\Lambda^\mu_\nu\frac{d}{dt}\Big[U(\mathbb{1},te^\nu)\Big]_{t=0}=\Lambda^\mu_\nu P^\nu.\tag{6.49}$$

□

6.2.4 Pure Lorentz transformation

Now we consider a pure Lorentz transformation $U(\Lambda)\equiv U(\Lambda,0)$, and we want to look at $U(\Lambda)|p,\sigma\rangle$. In order to see its decomposition into others $|k,\sigma'\rangle$, we apply a P^μ :

$$\begin{aligned}P^\mu U(\Lambda)|p,\sigma\rangle&=U(\Lambda)\Big(U(\Lambda)^{-1}P^\mu U(\Lambda)\Big)|p,\sigma\rangle\\&=U(\Lambda)(\Lambda^{-1})^\mu_\nu P^\nu|p,\sigma\rangle\\&=(\Lambda^{-1})^\mu_\nu p^\nu U(\Lambda)|p,\sigma\rangle.\end{aligned}\tag{6.50}$$

Thus the vector $U(\Lambda)|p,\sigma\rangle\in\mathscr{H}$ has $(\Lambda^{-1})^\mu_\nu p^\nu$ as eigenvalue for P^μ. If the p^μ's are seen as components of a 4-vector p, one can write

$$P^\mu U(\Lambda)|p,\sigma\rangle=(\Lambda p)^\mu U(\Lambda)|p,\sigma\rangle;$$

thus we naturally write

$$U(\Lambda)|p,\sigma\rangle=\sum_{\sigma'}C_{\sigma'\sigma}(\Lambda,p)|\Lambda p,\sigma'\rangle.\tag{6.51}$$

Note that we had not yet given anything about the nature of the p in the ket $|p,\sigma\rangle$ so we can *define* the product Λp by the fact that the ket $|\Lambda p,\sigma\rangle$ has eigenvalue $(\Lambda^{-1})^\mu_\nu p^\nu$ for the operator P^μ. So it is one of the $|p',\sigma'\rangle$.

6.2.5 Rebuilding of a basis for $\mathscr{H}$

From general considerations about the Lorentz group (many physicists had written very better books than me about) anyone knows that the only functions of the p^μ's which are invariant under all the Lorentz transformations are $p^2=\eta_\mu nup^\mu p^\nu$ and the sign of p^0 when $p^2<0$.

For any value of p^2 and sign of p^0, one consider a "standard vector" k. For example :

$$k=(1,0,0,1)\qquad\text{for } p^2=0,\tag{6.52a}$$

$$k=(1,0,0,0)\qquad\text{for } p^2<0,\ p^0>0,\tag{6.52b}$$

$$k=(-1,0,0,0)\qquad\text{for } p^2<0,\ p^0<0.\tag{6.52c}$$

With this convention, p can be written as $p=L(p)k$ for a suitable Lorentz transformation $L(p)$. The vector $U(L(p))|k,\sigma\rangle$ has eigenvalue $L(p)k$ for the operator P, thus it is a linear combination of some $|p,\sigma'\rangle$.

Now we will cheat and redefine our basis of the Hilbert space $\mathscr{H}$. First, we consider a fixed k; in other words, we build the state space for a given particle which has given momentum p. The basis vectors must be eigenvectors for the fours operators P^μ. As far as we say no more, any eigenvalue is possible. Thus our basis must be labelled by at least an element p of $\mathbb{R}^4$ with only one constraint: the value of p^2 (plus eventually the sign of p^0). So we define the $|k,\sigma\rangle$ to be such that

$$P^\mu|k,\sigma\rangle=k^\mu|k,\sigma\rangle.$$

Since we know that with this definition of $|k,\sigma\rangle$, the eigenvalue of $U(L(p))|k,\sigma\rangle$ for P^μ is p^μ, we *define* $|p,\sigma\rangle$ as

$$|p,\sigma\rangle=N(p)U(L(p))|k,\sigma\rangle.\tag{6.53}$$

where $N(p)$ is a normalization to be discussed later. With this construction, we have an eigenvector for any possible eigenvalue for P^μ. We have to show that these vectors are linearly independent.

The set of the $|p,\sigma\rangle$ with different p is free in $\mathscr{H}$ because they are eigenvectors for different eigenvalue of an hermitian operator [2] . There are no reason to think that the set of operators P^μ is complete; in other words, it remains not clear that there exist only one way to diagonalise the all the P^μ. The function of the extra label σ is to label different linearly independent vectors with same eigenvalue for P.

From now, we are interested in $|k,\sigma\rangle$ and $N(p)$.

6.2.6 Little group

We have :

$$\begin{aligned} U(\Lambda)|p,\sigma\rangle &= N(p)U(\Lambda L(p))|k,\sigma\rangle \\ &= N(p)U(L(\Lambda p))U\big(L(\Lambda p)^{-1}\Lambda L(p)\big)|k,\sigma\rangle, \end{aligned} \tag{6.54}$$

So we will try to understand the operation $L(\Lambda p)^{-1}\Lambda L(p)$. First remark that

$$U(L(\Lambda p)^{-1})|\Lambda p,\sigma\rangle = N(\Lambda p)|k,\sigma\rangle,$$

and then compute :

$$\begin{aligned} U(L(\Lambda p)^{-1}\Lambda L(p))N(p)|k,0\rangle &= U(L(\Lambda p)^{-1}\Lambda)|p,\sigma\rangle \\ &= U(L(\Lambda p)^{-1})\sum_{\sigma'} C_{\sigma'\sigma}(\Lambda,p)|\Lambda p,\sigma'\rangle \\ &= \sum_{\sigma'} C_{\sigma'\sigma}(\Lambda,p)N(\Lambda p)|k,\sigma'\rangle. \end{aligned} \tag{6.55}$$

The **little group** is the subgroup of the Lorentz transformations which leaves the chosen standard vector k invariant: $Wk = k$. For any W in the little group,

$$U(W)|k,\sigma\rangle = \sum_{\sigma'} D_{\sigma'\sigma}(W)|k,\sigma'\rangle$$

With this definition, the D's form a representation of the little group. Indeed for any V,W in the little group,

$$\begin{aligned} \sum_{\sigma'} D_{\sigma'\sigma}(VW)|k,\sigma'\rangle &= U(VW)|k,\sigma\rangle \\ &= U(V)\sum_{\sigma''} D_{\sigma''\sigma}(W)|k,\sigma''\rangle \\ &= \sum_{\sigma'\sigma''} D_{\sigma'\sigma''}(V)D_{\sigma''\sigma}(W)|k,\sigma'\rangle. \end{aligned} \tag{6.56}$$

Since we want the $|p,\sigma\rangle$ with different p and σ to form a basis of $\mathscr{H}$, they are linearly independent, then we can get rid of the sum over the σ' and keep the equation

$$D_{\sigma'\sigma}(VW) = \sum_{\sigma''} D_{\sigma'\sigma''}(VW)D_{\sigma''\sigma}(VW);$$

if we adopt a more “matricial” notation,

$$D(VW) = D(V)D(W). \tag{6.57}$$

We are now able to perform a step in the study of the vector $U(\Lambda)|p,\sigma\rangle$. We naturally define $W(\Lambda,p) = L(\Lambda p)^{-1}\Lambda L(p)$. This belongs to the little group[3]. Then,

$$\begin{aligned} U(\Lambda)|p,\sigma\rangle &= N(p)U(L(\Lambda p))U(W(\Lambda,p))|k,\sigma\rangle \\ &= N(p)\sum_{\sigma'} D_{\sigma'\sigma}(W)U(L(\Lambda p))|k,\sigma'\rangle \\ &= \frac{N(p)}{N(\Lambda p)}\sum_{\sigma'} D_{\sigma'\sigma}(W(\Lambda,p))|\Lambda p,\sigma'\rangle. \end{aligned} \tag{6.58}$$

But we have no constraint on the D's: it must just form a representation of the little group. Consequently, we are at a point in which our axioms are no more sufficient to continue the building of quantum field theory: we will get as many theories as representations of the little group.

[2] I did not checked that it is sufficient

[3] Pay attention that $L(p)$ depends implicitly on the choice of k.

The physical interpretation is the following : each type of particle has its own representation. When we consider a Hilbert space on which $U(\Lambda)$ acts via one given representation of the little group, we consider the Hilbert space which describes the corresponding particle. Note that the little group depends on the choice of k, and therefore depends on the particle which is studied (massive or not).

In this sense, a particle is a representation of the Poincaré group [4] . In particular, the nature of the index σ can change from the one representation to the other.

Remark 6.10.
As far as normalization is concerned, we will pose

$$N(p) = \sqrt{k^0/p^0}.$$

There are some good reasons to take it; but it is irrelevant from our group point of view of the theory.

6.2.7 Positive mass

This is the easy case. The choice of standard momentum is $k = \begin{pmatrix}1 & 0 & 0 & 0\end{pmatrix}$. One could believe that the little group is SO(3). It would be the case if we had chosen $L_+^\uparrow$ instead of $\mathrm{SL}(2,\mathbb{C})$ –see point 6.2.1. In our hermitian representation of $\mathbb{R}^4$, $k = \mathbb{1}$. Then a matrix of $\mathrm{SL}(2,\mathbb{C})$ which leaves it invariant fulfills

$$\lambda k \lambda^\dagger = \lambda\lambda^\dagger = \mathbb{1},$$

this is $\lambda \in SU(2)$. By the way, note that $\mathrm{SO}(3) = SU(2)/\mathbb{Z}_2$.

The celebrated "law of transformation" of a massive particle of spin j (integer or half integer) under the Lorentz transformation Λ is

$$U(\Lambda)|p,\sigma\rangle = \sqrt{\frac{(\Lambda p)^0}{p^0}} \sum_{\sigma'} D^{(j)}_{\sigma'\sigma}(W(\Lambda,p))|\Lambda p,\sigma'\rangle \tag{6.59}$$

where σ runs from $-j$ to j by step of 1.

6.2.8 Null mass

In the case of a null mass, the standard vector is $k = (1,0,0,1)$ and an element of the little group fulfils $Wk = k$. As the little group is part of the Lorentz group, this is an isometry, so

$$\langle Wt|Wk\rangle = \langle t|k\rangle \tag{6.60a}$$
$$\langle Wt|Wt\rangle = \langle t|t\rangle, \tag{6.60b}$$

for any $t \in \mathbb{R}^4$. Taking in particular $t = (1,0,0,0)$,

$$(Wt)^\mu k_\mu = t^\mu k_\mu = -1 \tag{6.61a}$$
$$(Wt)^\mu (Wt)_\mu = t^\mu t_\mu = -1. \tag{6.61b}$$

If we write $Wt = (a,b,c,d)$, the first relation gives $d = a-1$, so that $Wt = (1+\xi,\alpha,\beta,\xi)$, while the second one gives $\xi = (\alpha^2+\beta^2)/2$. The conclusion is that W acts on t as a certain Lorentz transformation $S(\alpha,\beta)$:

$$Wt = \begin{pmatrix}1+\xi\\ \alpha\\ \beta\\ \xi\end{pmatrix} = \begin{pmatrix}1+\xi & -\xi & \alpha & \beta\\ \alpha & -\alpha & 1 & 0\\ \beta & -\beta & 0 & 1\\ \xi & (1+\xi) & \alpha & \beta.\end{pmatrix}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}. \tag{6.62}$$

Be careful: it doesn't means that $W = S$, but $Wt = St$. However it is an information: $S(\alpha,\beta)^{-1}W$ is a Lorentz transformation which leaves t invariant. Then it is a spatial rotation. More precisely, since W and S conserve $(1,0,0,1)$, it is a rotation around the z axis: $S(\alpha,\beta)^{-1}W = R(\theta)$, and

$$W(\theta,\alpha,\beta) = S(\alpha,\beta)R(\theta) \tag{6.63}$$

is the most general element of the non massive little group.

[4] I think that the irreducibility of a representation is related to *elementary* particles.

Chapter 7

Relativistic fields and fibre bundle formalism

This chapter actually don't deal with *quantum* field theory in the sense that our wave functions aren't operators which acting on a Fock space. So this is just relativistic field theory.

7.1 Connections

7.1.1 Gauge potentials

Let us consider a **section** σ_α of P over $\mathcal{U}_\alpha$. It is a map $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$ such that $\pi \circ \sigma_\alpha = \text{id}$. A **connection** on P is a 1-form $\omega\colon T_pP \to \mathcal{G} \in \Omega^1(P)$ which satisfies the following two conditions:

$$\omega_p(Y_p^*) = Y, \tag{7.1a}$$

$$\omega(dR_g\xi) = g^{-1}\omega(\xi)g. \tag{7.1b}$$

The **gauge potential** of ω with respect of the local section σ_α is the 1-form on $\mathcal{U}_\alpha$ given by

$$A_\alpha(x)(v) = (\sigma_\alpha^*\omega)_x(v). \tag{7.2}$$

We will not always explicitly write the dependence of A_α in x. Now we consider another section $\sigma_\beta\colon \mathcal{U}_\beta \to P$ which is related on $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ to σ_α by $\sigma_\beta(x) = \sigma_\alpha(x) \cdot g_{\alpha\beta}(x)$ for a well defined map $g_{\alpha\beta}\colon \mathcal{U}_\alpha \cap \mathcal{U}_\beta \to G$.

Proposition 7.1.
The gauge potentials A_α and A_β are related by

$$A_\beta = g^{-1}A_\alpha g - g^{-1}dg. \tag{7.3}$$

Proof. By definition, for $v \in T_x\mathcal{U}_\alpha$,

$$A_\beta(v) = (\sigma_\beta^*\omega)_x(v) = \omega_{\sigma_\alpha(x)\cdot g_{\alpha\beta}(x)}\big((d\sigma_\beta)_x(v)\big).$$

We begin by computing $d\sigma_\beta(v)$. Let us take a path $v(t)$ in $\mathcal{U}_\alpha$ such that $v(0) = x$ and $v'(0) = v$. We have :

$$\begin{aligned}
(d\sigma_\beta)_x(v) &= \frac{d}{dt}\sigma_\beta(v(t))\Big|_{t=0}\\
&= \frac{d}{dt}\sigma_\alpha(v(t)) \cdot g_{\alpha\beta}(v(t))\Big|_{t=0}\\
&= \frac{d}{dt}\Big[\sigma_\alpha(v(t)) \cdot g_{\alpha\beta}(x)\Big]_{t=0} + \frac{d}{dt}\Big[\sigma_\alpha(x) \cdot g_{\alpha\beta}(v(t))\Big]_{t=0}\\
&= dR_{g_{\alpha\beta}(x)}d\sigma_\alpha(v) + \frac{d}{dt}\Big[\sigma_\alpha(x) \cdot g_{\alpha\beta}(x)e^{-ts}\Big]_{t=0}\\
&= dR_{g_{\alpha\beta}(x)}d\sigma_\alpha(v) + s^*_{\sigma_\alpha(x)\cdot g_{\alpha\beta}(x)}
\end{aligned} \tag{7.4}$$

where s is defined by the requirement that $g_{\alpha\beta}(x)^{-1}g_{\alpha\beta}(v(t))$ can be replaced in the derivative by e^{-ts}, so that we can replace $g_{\alpha\beta}(v(t))$ by $g_{\alpha\beta}(x)e^{-ts}$. As far as the derivatives are concerned, $e^{-ts} = g_{\alpha\beta}(x)^{-1}g_{\alpha\beta}(v(t))$, then

$$s = -\frac{d}{dt}g_{\alpha\beta}(x)^{-1}g_{\alpha\beta}(v(t))\Big|_{t=0} = -g_{\alpha\beta}(x)^{-1}dg_{\alpha\beta}(v),$$

this equality being a notation. Now, properties (7.1a) and (7.1b) make that

$$A_\beta(v) = g_{\alpha\beta}(x)^{-1}\omega_{\sigma_\alpha(x)}(d\sigma_\alpha(v))g_{\alpha\beta}(x) + s.$$

The thesis is just the same, with "reduced" notations (see section 7.4.2). . □

An explicit form for this transformation law is :

$$A_\beta(v) = \frac{d}{dt}\Big[g^{-1}e^{tA_\alpha(v)}g\Big]_{t=0} - \frac{d}{dt}\Big[g^{-1}g_{\alpha\beta}(v(t))\Big]_{t=0}, \tag{7.5}$$

where $g := g_{\alpha\beta}(x)$.

7.1.2 Covariant derivative

When we have a connection on a principal bundle, we can define a covariant derivative on any associated bundle. Let us quickly review it. An associated bundle is the semi-product $E = P \times_\rho V$ where V is a vector space on which acts the representation ρ of G. We denote the canonical projection by $\pi_p \colon E \to M$. The classes are taken with respect to the equivalence relation $(p, v) \sim (p \cdot g, \rho(g^{-1})v)$.

A **section** of E is a map $\psi \colon M \to E$ such that $\pi \circ \psi = \mathrm{id}$. We denote by $\Gamma(E)$ the set of all the sections of E. A section of E defines (and is defined by) an equivariant function $\hat\psi \colon P \to V$ such that

$$\psi(\pi(\xi)) = [\xi, \hat\psi(\xi)], \tag{7.6a}$$

$$\hat\psi(\xi \cdot g) = \rho(g^{-1})\hat\psi(\xi). \tag{7.6b}$$

For a section $\psi \in \Gamma(E)$, we define $\psi_{(\alpha)} \colon \mathcal{U}_\alpha \to V$ by

$$\psi_{(\alpha)}(x) = \hat\psi(\sigma(x)).$$

We saw in (4.118) that a covariant derivative on E is given by

$$(D_X\psi)_{(\alpha)}(x) = X_x\psi_{(\alpha)} - \rho_*\Big((\sigma_\alpha^*\omega)_x(X_x)\Big)\psi_{(\alpha)}(x). \tag{7.7}$$

Since $(d\psi)(X) = X(\psi)$, we can rewrite this formula in a simpler manner by forgetting the index α and the mention of X :

$$D\psi = d\psi - (\rho_* A_\alpha)\psi.$$

If we note $(\rho_* A_\alpha)\psi$ by $A_\alpha\psi$, we have

$$D\psi = d\psi - A\psi. \tag{7.8}$$

One has to understand that equation as a "notational trick" for (7.7). By the way, remark that $(\rho_* A_\alpha)$ is the only "reasonable" way for A to act on ψ.

7.2 Gauge transformation

A **gauge transformation** of a G-principal bundle is a diffeomorphism $\varphi \colon P \to P$ which satisfies

$$\pi \circ \varphi = \pi, \tag{7.9a}$$

$$\varphi(\xi \cdot g) = \varphi(\xi) \cdot g. \tag{7.9b}$$

In local coordinates, it can be expressed in terms of a function $\tilde\varphi_\alpha \colon \mathcal{U}_\alpha \to G$ by the requirement that

$$\varphi(\sigma_\alpha(x)) = \sigma_\alpha(x) \cdot \tilde\varphi_\alpha(x). \tag{7.10}$$

We have shown in proposition 4.33 that, if ω is a connection 1-form on P, the form $\varphi \cdot \omega := \varphi^*\omega$ is still a connection 1-form on P. Of course, with the same section σ_α than before, we can define a gauge potential $(\varphi \cdot A)_\alpha$ for this new connection. We will see how it is related to A_α. The reader can guess the result (it will be the same as proposition 7.1). We show it.

Proposition 7.2.

$$(\varphi \cdot A) = \tilde\varphi^{-1}A\tilde\varphi - \tilde\varphi^{-1}d\tilde\varphi. \tag{7.11}$$

Proof. Let us consider $x \in \mathcal{U}_\alpha$, and $v \in T_x\mathcal{U}_\alpha$, the vector which is tangent to the curve $v(t) \in \mathcal{U}_\alpha$. We compute

$$\sigma_\alpha^*(\varphi^*\omega)_x(v) = \omega_{(\varphi\circ\sigma_\alpha)(x)}((d\varphi \circ d\sigma_\alpha)(v)),$$

but equation (7.10) makes

$$\begin{aligned}(d\varphi \circ d\sigma_\alpha)(v) &= \frac{d}{dt}\varphi(\sigma_\alpha(v(t)))\Big|_{t=0} \\ &= \frac{d}{dt}\sigma_\alpha(v(t)) \cdot \tilde{\varphi}_\alpha(v(t))\Big|_{t=0}.\end{aligned} \tag{7.12}$$

Now, we are in the same situation as in equation (7.4). □

If $\psi\colon M \to E$ is a section of E, the gauge transformation $\varphi\colon P \to P$ acts on ψ by

$$\widehat{\varphi\cdot\psi}(\xi) = \hat{\psi}(\varphi^{-1}(\xi)). \tag{7.13}$$

On the other hand, φ acts on the covariant derivative (and the potential) : $\varphi \cdot D$ is the covariant derivative of the connection $\varphi \cdot \omega$. Of course, we define

$$(\varphi \cdot D)\psi = d\psi - (\varphi \cdot A)\psi. \tag{7.14}$$

Lemma 7.3.
If $\varphi\colon P \to P$ is a gauge transformation, then

(i) φ^{-1} *is also a gauge transformation and* $(\widetilde{\varphi^{-1}})_\alpha(x) = \tilde{\varphi}_\alpha(x)^{-1}$,

(ii) $(\varphi \cdot \psi)_{(\alpha)}(x) = \rho(\tilde{\varphi}_x^{-1})\psi_{(\alpha)}(x)$.

Proof. The first part is clear while the second is a computation :

$$(\varphi\cdot\psi)_{(\alpha)} = \widehat{\varphi\cdot\psi}(\sigma_\alpha(x)) = \hat{\psi}(\varphi^{-1}(\sigma_\alpha(x))) = \hat{\psi}(\sigma_\alpha(x)\cdot\tilde{\varphi}_\alpha(x)^{-1}) = \rho(\tilde{\varphi}_\alpha(x))\psi_{(\alpha)}(x). \tag{7.15}$$

□

Now, we will proof the main theorem: the one which explains why the covariant derivative is "covariant".

Theorem 7.4.
The covariant derivative D fulfils a "covariant" transformation rule under gauge transformations:

$$(\varphi \cdot D)(\varphi^{-1}\cdot\psi) = \varphi^{-1}(D\psi). \tag{7.16}$$

Remark 7.5.
Let us use more intuitive notations: we write (7.11) under the form $A' = g^{-1}Ag - g^{-1}dg$. If we have two sections ψ and ψ', they are necessarily related by a gauge transformation: $\psi' = g^{-1}\psi$. Then, the theorem tells us that the equation $D\psi = d\psi - A\psi$ becomes $D'\psi' = g^{-1}D\psi$ "under a gauge transformation". This is: $D\psi$ transforms under a gauge transformation as $d\psi$ transforms under a constant linear transformation. This is the reason why D is a covariant *derivative. The physicist way to write (7.16) is*

$$D'\psi' = g^{-1}D\psi \tag{7.17}$$

Proof of theorem 7.4. First, we look at $(\varphi \cdot A)\psi_\alpha$. Using all the notational tricks used to give a sens to $A\psi$, we write :

$$[(\varphi\cdot A)_X\psi]_{(\alpha)}(x) = (\varphi\cdot A)_X\psi_{(\alpha)}(x) = \rho_*(\varphi\cdot A(X))\psi_{(\alpha)}(x).$$

But we know that $\varphi \cdot A = \tilde{\varphi}^{-1}A\tilde{\varphi} - \tilde{\varphi}^{-1}d\tilde{\varphi}$, then

$$\begin{aligned}(\varphi\cdot A)_X\psi_{(\alpha)}(x) &= \rho_*(\tilde{\varphi}^{-1}A(X)\tilde{\varphi})\psi_{(\alpha)}(x) \\ &\quad - \rho_*(\tilde{\varphi}^{-1}d\tilde{\varphi}(X))\psi_{(\alpha)}(x) \\ &= \frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}e^{tA(X)}\tilde{\varphi})\psi_{(\alpha)}(x)\Big]_{t=0} \\ &\quad - \frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}\tilde{\varphi}(X_t))\psi_{(\alpha)}(x)\Big]_{t=0}\end{aligned} \tag{7.18}$$

Now, we have to write this equation with $\varphi^{-1}\cdot\psi$ instead of ψ. Using lemma 7.3, we find :

$$\begin{aligned}(\varphi\cdot A)_X(\varphi^{-1}\cdot\psi)_{(\alpha)}(x) = \frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}e^{tA(X)}\tilde{\varphi}\tilde{\varphi}^{-1})\psi_{(\alpha)}(x)\Big]_{t=0}\\ -\frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}\tilde{\varphi}(X_t)\tilde{\varphi}^{-1})\psi_{(\alpha)}(x)\Big]_{t=0}\end{aligned} \tag{7.19}$$

After simplification, the first term is a term of the thesis: $\tilde{\varphi}(x)^{-1}(A\psi)_\alpha(x)$ and we let the second one as it is. Now, we turn our attention to the second term of (7.16); the same argument gives:

$$\begin{aligned}d(\varphi^{-1}\psi_{(\alpha)})_xX &= \frac{d}{dt}\Big[(\varphi^{-1}\cdot\psi)_{(\alpha)}(X_t)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\rho(\tilde{\varphi}(X_t)^{-1})\psi_{(\alpha)}(X_t)\Big]_{t=0}\\ &= \frac{d}{dt}\Big[\rho(\tilde{\varphi}(X_t)^{-1})\psi_{(\alpha)}(x)\Big]_{t=0}+\frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1})\psi_{(\alpha)}(X_t)\Big]_{t=0}.\end{aligned} \tag{7.20}$$

The second term is $\tilde{\varphi}^{-1}d\psi_\alpha(X)$. In definitive, we need to prove that the two exceeding terms cancel each other:

$$\frac{d}{dt}\Big[\rho(\tilde{\varphi}^{-1}\tilde{\varphi}(X_t)\tilde{\varphi}^{-1})\psi_{(\alpha)}(x)\Big]_{t=0}+\frac{d}{dt}\Big[\rho(\tilde{\varphi}(X_t)^{-1})\psi_{(\alpha)}(x)\Big]_{t=0} \tag{7.21}$$

must be zero.

One can find a $g(t)\in G$ such that $\tilde{\varphi}(X_t)=\tilde{\varphi}g(t)$, $g(0)=e$. Then, what we have in the ρ of these two terms is respectively $g(t)\tilde{\varphi}^{-1}$ and $g(t)^{-1}\tilde{\varphi}^{-1}$. As far as the derivative are concerned, $g(t)$ can be written as e^{tZ} for a certain $Z\in\mathcal{G}$. So we see that $g(t)^{-1}=e^{-tZ}$ and the derivative will come with the right sign to makes the sum zero. □

Remark 7.6.
If we naively make the computation with the notations of remark 7.5, we replace $\psi'=g^{-1}\psi$ and $A'=g^{-1}Ag-g^{-1}dg$ in

$$D'\psi'=d\psi'-A'\psi',$$

using some intuitive "Leibnitz formulas", we find : $D'\psi'=dg^{-1}\psi+g^{-1}d\psi+g^{-1}A\psi+g^{-1}dgg^{-1}\psi$. It is exactly $g^{-1}d\psi+g^{-1}A\psi$ with two additional terms: $dg^{-1}\psi$ and $g^{-1}dgg^{-1}\psi$. One sees that these are precisely the two terms of the expression (7.21). We will give a sens to this "naive" computation in section 7.4.2.

7.3 A bite of physics

7.3.1 Example: electromagnetism

Let us consider the electromagnetism as the simplest example of a gauge invariant physical theory. We first discuss the theory of free electromagnetic field (this is: without taking into account the interactions with particles) from Maxwell's equations, see [47, 57]. The electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ are subject to following relations:

$$\nabla\cdot\mathbf{E}=\rho, \tag{7.22a}$$
$$\nabla\cdot\mathbf{B}=0, \tag{7.22b}$$
$$\nabla\times\mathbf{E}+\partial_t\mathbf{B}=0, \tag{7.22c}$$
$$\nabla\times\mathbf{B}-\partial_t\mathbf{E}=\mathbf{j}. \tag{7.22d}$$

Comparing (7.22a) and (7.22b), we see that Maxwell's theory does not incorporate magnetic monopoles. Suppose that we can use the Poincaré lemma. Equation (7.22b) gives a vector field $\mathbf{A}$ such that $\mathbf{B}=\nabla\times\mathbf{A}$, so that (7.22c) becomes $\nabla\times(\mathbf{E}+\partial_t\mathbf{A})=0$ which gives a scalar field ϕ such that $-\nabla\cdot\phi=\mathbf{E}+\partial_t\mathbf{A}$.

Now the equations (7.22a)–(7.22d) are equations for the potentials $\mathbf{A}$ and ϕ, and we find back the "physical" field by

$$\mathbf{B}=\nabla\times\mathbf{A}, \tag{7.23a}$$
$$\mathbf{E}=-\nabla\phi-\partial_t\mathbf{A}. \tag{7.23b}$$

One can easily see that there are several choice of potentials which describe the same electromagnetic field. Indeed, if

$$\mathbf{A}'=\mathbf{A}+\nabla\lambda, \tag{7.24a}$$
$$\phi'=\phi-\partial_t\lambda, \tag{7.24b}$$

the electromagnetic field given (via (7.23)) by $\{\phi', \mathbf{A}'\}$ is the same as the one given by $\{\phi, \mathbf{A}\}$

The Maxwell's equations can be written in a more "covariant" way by defining

$$F = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ \cdot & 0 & -B_z & \cdot \\ \cdot & \cdot & 0 & -B_x \\ \cdot & -B_y & \cdot & 0 \end{pmatrix}, \tag{7.25}$$

$F^{\mu\nu} = -F^{\nu\mu}$ and

$$J = \begin{pmatrix} c\rho & j_x & j_y & j_z \end{pmatrix}.$$

We also define $\star F^{\alpha\beta} = \frac{1}{2}e^{\alpha\beta\lambda\mu}F_{\lambda\mu}$. With all that, Maxwell's equations read:

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= \mu_0 J^\nu, \\ \partial_\alpha \star F^{\alpha\beta} &= 0. \end{aligned} \tag{7.26}$$

If we define

$$A = \begin{pmatrix} \frac{\phi}{c} & -A_x & -A_y & -A_z \end{pmatrix}, \tag{7.27}$$

the physical fields are given by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu.$$

The **gauge invariance** of this theory is the fact that

$$F'_{\mu\nu} = \partial_\mu A'_\nu - \partial_\nu A'_\mu = F_{\mu\nu} \tag{7.28a}$$

when

$$A'_\mu(x) = A_\mu(x) + \partial_\mu f(x) \tag{7.28b}$$

for any scalar *function* f (to be compared with (7.24)).

This is: in the picture of the world in which we see the A as fundamental field of physics, several (as much as you have functions in $C^\infty(\mathbb{R}^4)$) fields A, A',... describe the *same* physical situation because the fields $\mathbf{E}$ and $\mathbf{B}$ which acts on the particle are the same for A and A'.

Now, we turn our attentions to the interacting field theory of electromagnetism. As far as we know, the electron makes interactions with the electromagnetic field via a term $\overline{\psi}A_\mu\psi$ in the Lagrangian. The free Lagrangian for an electron is

$$\mathcal{L} = \overline{\psi}(\gamma^\mu\partial_\mu + m)\psi. \tag{7.29}$$

The easiest way to include a $\overline{\psi}A\psi$ term is to change ∂_μ to $\partial_\mu + A_\mu$. But we want to preserve the powerful gauge invariance of classical electrodynamics, then we want the new Lagrangian to keep unchanged if we do

$$A_\mu \to A'_\mu = A_\mu - i\partial_\mu\phi. \tag{7.30}$$

In order to achieve it, we remark that the ψ must be transformed *simultaneously* into

$$\psi'(x) = e^{i\phi(x)}\psi(x). \tag{7.31}$$

The conclusion is that if one want to write down a Lagrangian for QED, one must find a Lagrangian which remains unchanged under certain transformation $A \to A'$ and $\psi \to \psi'$. In other words the set $\{\psi, A\}$ of fields which describe the world of an electron in an electromagnetic field is not well defined from data of the physical situation alone: it is defined up to a certain invariance which is naturally called a **gauge invariance**.

Remark 7.7.
In the physics books, the matter is presented in a slightly different way. We observe that the Lagrangian (7.29) is invariant under

$$\psi(x) \to \psi'(x) = e^{i\alpha}\psi(x) \tag{7.32}$$

for any constant α. *One can see that the associated conserved current (Noether) is closely related to the electric current. The idea (of Yang-Mills) is to develop this symmetry. Since the symmetry (7.32) depends only on a constant, we say it a* ***global*** *symmetry; we will simultaneously add a new field A_μ and upgrade (7.32) to a* ***local*** *symmetry:*

$$\psi(x) \to \psi'(x) = e^{i\phi(x)}\psi(x). \tag{7.33}$$

Then, we deduce the transformation law of A_μ.

Because of the form of (7.31), we say that the electromagnetism is a $U(1)$-gauge theory. The fact that this is an abelian group have a deep physical meaning and many consequences.

7.3.2 Little more general, slightly more formal

The aim of this text is to interpret the field A as a gauge potential for a connection. But equation (7.30) is not exactly the expected one which is (7.11). The point is that equation (7.30) concerns a theory in which the gauge transformation of the field was a simple multiplication by a scalar field, so that simplifications as $e^{-i\phi(x)}A_\mu(x)e^{i\phi(x)} = A_\mu(x)$ are allowed.

Now, we consider a vector space V, a manifold M and a function $\psi\colon M \to V$ which "equation of motion" is

$$L^i(\partial_i + m_i)\psi = 0$$

Where we imply an unit matrix behind ∂ and m; the indices i, j are the (local) coordinates in M and a, b, the coordinates in V. Let G be a matrix group which acts on V. If ψ is a solution, $\Lambda^{-1}\psi$ is also a solution as far as Λ is a constant –does not depend on $x \in M$– matrix of G. In other words, $L^i(\partial_i + m_i)\psi_a = 0$ for all a implies $L^i(\partial_i + m_i)((\Lambda^{-1})^b_a\psi_b) = 0$.

The function, $\psi'(x) = \Lambda(x)^{-1}\psi(x)$ is no more a solution. If we want it to be solution of the same equation as ψ, we have to change the equation and consider

$$L^i(\partial_i + A_i + m_i)\psi = 0.$$

This equation is preserved under the *simultaneous* change

$$\begin{cases} \psi'_a = (\Lambda^{-1})^b_a\psi_b \\ (A'_i)^a_b = (\Lambda^{-1})^c_b(A_i)^d_c(\Lambda^a_d) - (\partial_i\Lambda^{-1})^d_b\Lambda^a_d. \end{cases} \tag{7.34}$$

The second line show that the formalism in which A is a connection is the good one to write down covariant equations. This has to be compared with (7.3). Logically, a theory which includes an invariance under transformations as (7.34) is called a G-gauge theory.

7.3.3 A "final" formalism

Now, we work with fields which are sections of some fiber bundle build over M, the physical space. More precisely, let G be a matrix group.

Problem and misunderstanding 30.
For sure, it also works for a much lager class of groups. Which one ?

We search for a theory which is "locally invariant under G". In order to achieve it, we consider a G-principal bundle P over M and the associated bundle $E = P \times_\rho V$ for a certain vector space V, and a representation ρ of G on V. Typically, V is $\mathbb{C}$ or the vector space on which the spinor representation acts.

The physical fields are sections $\psi\colon M \to E$. If we choose some reference sections $\sigma_\alpha\colon M \to P$, they can be expressed by $\psi_{(\alpha)}(x) = \hat\psi(\sigma_\alpha(x))$. We translate the idea of a local invariance under G by requiring an invariance under

$$\psi'_{(\alpha)}(x) = \rho(g(x))\psi_{(\alpha)}(x)$$

for every $g\colon M \to G$. By (ii) of lemma 7.3, we see that $\psi'_{(\alpha)}(x) = (\varphi^{-1}\cdot\psi)_{(\alpha)}(x)$, where $\varphi\colon P \to P$ is the gauge transformation given by

$$\varphi(\sigma_\alpha(x)) = \sigma_\alpha(x)\cdot g(x).$$

We want ψ and ψ' to "describe the same physics". From a mathematical point of view, we want ψ and ψ' to *satisfy the same equation.* It is clear that equation $d\psi = 0$ will not work.

The trick is to consider any connection ω on P and the gauge potential A of ω. In this case the equation

$$(d - A)\psi = 0 \qquad \text{or} \qquad D\psi = 0 \tag{7.35}$$

is preserved under

$$\begin{aligned} A &\to \varphi\cdot A, \\ \psi &\to \varphi^{-1}\cdot\psi. \end{aligned}$$

Theorem 7.4 powa !

In this sense, we say that equation (7.35) is gauge invariant, and is thus taken by physicists to build some theories when they need a "local G-covariance". This gives rise to the famous Yang-Mills theories.

In this picture the matter field ψ and the bosonic field A are both defined from a $U(1)$-principal bundle. When physicists say

ψ transforms as "blahblah" under a $U(1)$ transformation,

they mean that ψ is a section of an $U(1)$-associated bundle; when they say

A transforms as "blahblah" under a $U(1)$ transformation,

they mean that A is the gauge potential of a connection on a $U(1)$-principal bundle. In each case, the "blahblah" denotes an irreducible[1] representation of $U(1)$.

Remark 7.8.
The mathematics of equation (7.35) only requires a $\mathcal{G}$-valued connection on P. There are several physical constraints on the choice of the connection. These give rise to interaction terms between the gauge bosons. We will not discuss it at all. This a matter of books about quantum field theories.

The most used Yang-Mills groups in physics are $U(1)$ for the QED, $SU(2)$ for the weak interactions and $SU(3)$ for chromodynamic.

7.4 Curvature

7.4.1 Intuitive setting

From the $\mathcal{G}$-valued connection 1-form ω on P, we may define its **curvature 2-form** :

$$\Omega = d\omega + \omega \wedge \omega. \tag{7.36}$$

As before, we can see Ω as a 2-form on M instead of P. For this, we just need some sections $\sigma_\alpha \colon \mathcal{U}_\alpha \to P$ and define

$$F_\alpha = \sigma_\alpha^* \Omega. \tag{7.37}$$

This F is called the **Yang-Mills field strength**. The question is now to see how does it transform under a change of chart ? What is $F_\beta = \sigma_\beta^* \Omega$ in terms of F_α ?

Theorem 7.9.

$$F_\beta = g^{-1} F_\alpha g. \tag{7.38}$$

Naive proof. Let us accept $F_\beta = dA_\beta + A_\beta \wedge A_\beta$. With proposition 7.1, we can perform a simple computation with all the intuitive "Leibnitz rules" :

$$dA_\beta = -g^{-1} dg\, g^{-1} \wedge A_\alpha g + g^{-1} dA_\alpha g + g^{-1} A_\alpha \wedge dg - g^{-1} dg\, g^{-1} \wedge dg,$$

and

$$A_\beta \wedge A_\beta = g^{-1} A_\alpha g \wedge g^{-1} A_\alpha g + g^{-1} A_\alpha g \wedge g^{-1} dg + g^{-1} dg \wedge g^{-1} A_\alpha g + g^{-1} dg \wedge g^{-1} dg.$$

The sum is obviously the announced result. □

This proof seems too beautiful to be false[2]. We will now try to give a sense to this computation. A complete proof of the theorem is reported until page 254.

First, note that we can't try to find a relation like $d(g\omega) = dg \wedge \omega + g\, d\omega$. Pose $A_x = g(x)\omega_x$:

$$A_x(v) = \frac{d}{dt} g(x) e^{t\omega_x(v)} \Big|_{t=0}.$$

Using

$$(d\alpha)(v, w) = v(\alpha(w)) - w(\alpha(v)) - \alpha([v, w]),$$

we are led to write

$$w(A(v)) = d(A(v))w = \frac{d}{du} A_{w_u}(v) \Big|_{u=0} = \frac{d}{du} \frac{d}{dt} \Big[g(w_u) e^{t\omega_{w_u}(v)} \Big]_{t=0} \Big|_{u=0}. \tag{7.39}$$

But at $t = u = 0$, the expression in the bracket is $g(x)$, and not e. Then the whole expression is not an element of $\mathcal{G}$. In other words, the problem is that for $g \colon M \to G$, we have $dg_x \colon T_x M \to T_{g(x)} G \neq T_e G$.

Now, remark that in our matter, the problem will not arise because in the expressions $A_\beta = g^{-1} A_\alpha g + g^{-1} dg$, each term has a g and a g^{-1}.

[1] Irreducibility is for elementary particles
[2] More precisely, it is as beautiful as we want it to be true.

Lemma 7.10.

$$d(g^{-1})_x(v) = -g(x)^{-1}dg(v)g(x)^{-1}. \tag{7.40}$$

Proof. Let v_t be a path which defines the vector v, and define $Y \in \mathcal{G}$ such that as far as the derivative are concerned, we have $g(v_t) = g(x)e^{tY}$. Then,

$$g(g^{-1})(v) = \frac{d}{dt}\Big[g(v_t)^{-1}\Big]_{t=0} = \frac{d}{dt}\Big[e^{-tY}g(x)^{-1}\Big]_{t=0}.$$

But on the other hand,

$$g^{-1}dg(v)g^{-1} = \frac{d}{dt}\Big[g(x)^{-1}g(v_t)g(x)^{-1}\Big]_{t=0} = \frac{d}{dt}\Big[e^{tY}g(x)^{-1}\Big]_{t=0},$$

thus $d(g^{-1})_x(v) = -g(x)^{-1}dg(v)g(x)^{-1}$, as we want. □

7.4.2 A digression: $T_Y\mathcal{G}$ and $\mathcal{G}$

We define two product: $G \times \mathcal{G} \to TG$ and $\mathcal{G}\times\mathcal{G} \to \mathcal{G}$. If $g \in G$ and $X \in \mathcal{G}$, we put

$$gX = \frac{d}{dt}ge^{tX}\Big|_{t=0}, \tag{7.41a}$$

and if X, $Y \in \mathcal{G}$,

$$XY = \frac{d}{dt}\frac{d}{du}\Big[e^{tX}e^{uY}\Big]_{\substack{u=0\\t=0}}. \tag{7.41b}$$

We naturally define the product of a $\mathcal{G}$-valued 1-form A by an element $g \in G$ by $(gA)v = gA(v)$.

Note that gX does not belong to $\mathcal{G}$ but to T_gG. Fortunately, in the expressions which we will meet, there will always be a g^{-1} to save the situation.

Let us now see a great consequence of the second definition.

Proposition 7.11.
The formula

$$XY - YX = [X, Y]. \tag{7.42}$$

links the formal product inside the Lie algebra and the Lie bracket.

In order to get a real proof (given from page 253) of this, we have to give some precisions about derivatives as (7.41b). We consider the expression

$$\frac{d}{du}\left(\frac{d}{dt}c_u(t)\Big|_{t=0}\right)_{u=0},$$

which will be more simply written as :

$$\frac{d}{du}\frac{d}{dt}\Big[c_u(t)\Big]_{\substack{t=0\\u=0}} \tag{7.43}$$

with $c_u(t) \in G$ for all u, t; $c_u(0) = e$ for all u and $c_0'(0) = Y \in \mathcal{G}$ where the prime stands for the derivative with respect of t. So $\frac{d}{dt}c_u(t)\big|_{t=0} \in \mathcal{G}$ for each u and (7.43) belongs to $T_Y\mathcal{G}$. But we know that $\mathcal{G}$ is a vector space, then $T_Y\mathcal{G} \simeq \mathcal{G}$, the isomorphism being given by the following idea: if $\{\partial_i\}$ is a basis of $\mathcal{G}$ and $\{e_i\}$ the corresponding basis of $T_Y\mathcal{G}$, we define the action of $A^ie_i \in T_Y\mathcal{G}$ on $f\colon G \to \mathbb{R}$ by $(A^ie_i)f := A^i\partial_i f$.

Lemma 7.12.
Seen as an equality in $\mathcal{G}$, for $f\colon G \to \mathbb{R}$ we have :

$$\frac{d}{du}\frac{d}{dt}\Big[c_u(t)\Big]_{\substack{t=0\\u=0}}f = \frac{d}{du}\frac{d}{dt}\Big[f(c_u(t))\Big]_{\substack{t=0\\u=0}}. \tag{7.44}$$

Proof. Let us consider $X_u = X_u^i\partial_i = c_u'(0)$ and $X_0 = Y$. We naturally have

$$X_uf = \frac{d}{dt}f(c_u(t))\Big|_{t=0}, \qquad \text{and} \qquad \frac{d}{du}X_u\Big|_{u=0} \in T_Y\mathcal{G}. \tag{7.45}$$

Now, we consider a function $h\colon \mathcal{G}\to\mathbb{R}$ and compute :

$$\frac{d}{du}\Big[X_u\Big]_{u=0}h=\frac{d}{du}\Big[h(X_u)\Big]_{u=0}=\frac{d}{du}h(\frac{d}{dt}\Big[c_u(t)\Big]_{t=0})\Big|_{u=0}.$$

If $\{\partial_i\}$ is a basis of $\mathcal{G}$ and $\{e_i\}$, the corresponding one of $T_Y\mathcal{G}$, thus

$$\frac{d}{du}\Big[X_u\Big]_{u=0}h=\frac{\partial h}{\partial e_i}\Big|_Y\frac{d}{du}\frac{d}{dt}\Big[c_u^i(t)\Big]_{\substack{t=0\\u=0}}. \tag{7.46}$$

So, we can write

$$\frac{d}{du}\Big[X_u\Big]_{u=0}=\frac{d}{du}\frac{d}{dt}\Big[c_u^i(t)\Big]_{\substack{t=0\\u=0}}\frac{\partial}{\partial e_i}\Big|_Y,$$

and, as element of $\mathcal{G}$, we consider

$$\frac{d}{du}\Big[X_u\Big]_{u=0}=\frac{d}{du}\frac{d}{dt}\Big[c_u^i(t)\Big]_{\substack{t=0\\u=0}}\partial_i|_e.$$

Now, we can compute the action of $\frac{d}{du}X_u\big|_{u=0}$ on a function $f\colon G\to\mathbb{R}$ as

$$\begin{aligned}\frac{d}{du}\Big[X_u\Big]_{u=0}f&=\frac{d}{du}\frac{d}{dt}\Big[c_u^i(t)\Big]_{\substack{t=0\\u=0}}\frac{\partial f}{\partial x^i}\Big|_e\\&=\frac{d}{du}\Big[\frac{\partial f}{\partial x^i}\Big|_e\frac{d}{dt}c_u^i(t)\Big|_{t=0}\Big]_{u=0}\\&=\frac{d}{du}\Big[\frac{d}{dt}f(c_u(t))\Big|_{t=0}\Big]_{u=0}.\end{aligned} \tag{7.47}$$

Problem and misunderstanding **31.**
Je ne sais pas pourquoi tout d'un coup la dernière équation était commentée, et donc la phrase n'était pas finie.

□

Proof of proposition 7.11. From this, we can precise our definition of XY when X, $Y\in\mathcal{G}$. The product XY acts on $f\colon G\to\mathbb{R}$ by

$$(XY)f=\frac{d}{dt}\frac{d}{du}\Big[f(e^{tX}e^{uY})\Big]_{\substack{u=0\\t=0}}.$$

We can get a more geometric interpretation of this. We know how to build a left invariant vector field $\tilde{Y}$ from any $Y\in\mathcal{G}$: for each $g\in G$ we just have to define

$$\tilde{Y}_g(f)=\frac{d}{ds}\Big[f(gY(s))\Big]_{s=0}.$$

First remark: $\tilde{Y}_g$ is precisely the object that previously named "gY". In order to construct the basis blocks of the formula $XY-YX=[X,Y]$, we turn our attention to $\tilde{X}_e\tilde{Y}$. It is clear that $\tilde{Y}(f)$ is a function from G to $\mathbb{R}$, so we can apply $\tilde{X}_e$ on it. If X_t is a path which gives the vector $\tilde{X}_e$ (for example: $X_t=e^{tX}$), we have

$$\tilde{X}_e(\tilde{Y}(f))=\frac{d}{dt}\Big[\tilde{Y}(f)_{X_t}\Big]_{t=0}=\frac{d}{du}\frac{d}{dt}\Big[f(X_tY(u))\Big]_{\substack{t=0\\u=0}}=\frac{d}{du}\frac{d}{dt}\Big[f(e^{tX}e^{uY})\Big]_{\substack{t=0\\u=0}}. \tag{7.48}$$

Thus we have: $XY=\tilde{X}_e\tilde{Y}$, but it is clear that $[\tilde{X},\tilde{Y}]_e=\tilde{X}_e\tilde{Y}-\tilde{Y}_e\tilde{X}$. The claim reads now: $[\tilde{X},\tilde{Y}]_e=[X,Y]$. We can actually take it as de *definition* of $[X,Y]$. It is done in [3]. The link with the definition in terms of successive derivations of $\mathbf{Ad}_g(x)=gxg^{-1}$ is not trivial but it can be done. □

Now, we can give a powerful definition of the wedge for two $\mathcal{G}$-valued 1-forms. If A, $B\in\Omega^1(M,\mathcal{G})$ and v, $w\in\mathfrak{X}(M)$, we define

$$(A\wedge B)(v,w)=A(v)B(w)-A(w)B(v). \tag{7.49}$$

For A^2, we find back the usual definition :

$$(A\wedge A)(v,w)=A(v)A(w)-A(w)A(v)=[A(v),A(w)].$$

One can see that for any section $\sigma_\alpha\colon\mathcal{U}_\alpha\to P$, we have

$$\sigma_\alpha^*(A\wedge A)=(\sigma_\alpha^*A)\wedge(\sigma_\alpha^*A). \tag{7.50}$$

Lemma 7.13.
If A and B are two $\mathcal{G}$-valued 1-forms, one can make "simplifications" as

$$(Ag) \wedge (g^{-1}B) = A \wedge B. \tag{7.51}$$

Proof. We just have to prove that for $A, B \in \mathcal{G}$, $(Ag)(g^{-1}B) = AB$ with definitions (7.41a) and (7.41b). Remark that $Ag = \frac{d}{ds}\Big[e^{sA}g\Big]_{s=0}$, so

$$e^{tAg} = \exp(t\,\frac{d}{ds}e^{sA}g\Big|_{s=0}) = \exp(\frac{d}{ds}e^{stA}g\Big|_{s=0}) = e^{tA}g.$$

Therefore

$$(Ag)(g^{-1}B) = \frac{d}{dt}\frac{d}{du}\Big[e^{tAg}e^{ug^{-1}B}\Big]_{\substack{u=0\\t=0}} = \frac{d}{dt}\frac{d}{du}\Big[e^{tA}gg^{-1}e^{uB}\Big]_{\substack{u=0\\t=0}} = AB.$$

□

Lemma 7.14.

$$F_\beta = dA_\beta + A_\beta^2. \tag{7.52}$$

Proof. This is a direct consequence of (7.50) and $[\sigma_\beta^*, d] = 0$. □

Now, we can prove the theorem.

Ultimate proof of theorem 7.9. First we compute $d(g^{-1}A_\alpha g)$. In order to do this, remark that the 1-form $g^{-1}A_\alpha g$ is explicitly given on $v \in \mathfrak{X}(M)$ by

$$(g^{-1}A_\alpha g)(v)_x = \frac{d}{dt}\Big[g(x)^{-1}e^{tA(v)_x}g(x)\Big]_{t=0}.$$

For all $x \in M$, this expression is an element of $\mathcal{G}$; then we can say that $(g^{-1}A_\alpha g)(v)$ is a map $(g^{-1}A_\alpha g)(v)\colon M \to \mathcal{G}$. So it is unambiguous to write $w((g^{-1}A_\alpha g)(v)) \in \mathcal{G}$ for $w \in T_xM$.

We will use the formula

$$d(g^{-1}A_\alpha g)(v,w) = v(g^{-1}A_\alpha g)(w) - w(g^{-1}A_\alpha g)(v) - (g^{-1}A_\alpha g)([v,w]).$$

As $w((g^{-1}A_\alpha g)(v)) = d((g^{-1}A_\alpha g)(v))w$, we have

$$\begin{aligned}
w((g^{-1}A_\alpha g)(v)) &= \frac{d}{du}(g^{-1}A_\alpha g)(v)_{w_u}\Big|_{u=0}\\
&= \frac{d}{du}\frac{d}{dt}\Big[g(w_u)^{-1}e^{tA(v)_{w_u}}g(w_u)\Big]_{t=0}\Big|_{u=0}\\
&= \frac{d}{dt}\frac{d}{du}\Big[g(w_u)^{-1}\Big]_{u=0}e^{tA(v)_x}g(x)\Big|_{t=0}\\
&\quad+\frac{d}{dt}g(x)^{-1}\frac{d}{du}\Big[e^{tA(v)_{w_u}}\Big]_{u=0}g(x)\Big|_{t=0}\\
&\quad+\frac{d}{dt}g(x)^{-1}e^{tA(v)_x}\frac{d}{du}\Big[g(w_u)\Big]_{u=0}\Big|_{t=0}\\
&= d(g^{-1})(w)A(v)_xg(x)\\
&\quad+g(x)^{-1}w_x(A(v))g(x)\\
&\quad+g(x)^{-1}A(v)_xdg(w).
\end{aligned} \tag{7.53}$$

On the other hand, one easily finds that

$$(g^{-1}A_\alpha g)([v,w]) = g(x)^{-1}A([v,w])g(x).$$

Using lemma 7.10, we have

$$\begin{aligned}
d(g^{-1}A_\alpha g)_x(v,w) = &-g(x)^{-1}dg(v)g(x)^{-1}A(w)_xg(x) + g(x)^{-1}v(A(w))g(x)\\
&+g(x)^{-1}A(w)_xdg(v)_x\\
&+g(x)^{-1}dg(w)_xg(x)^{-1}A(v)_xg(x) - g(x)^{-1}w(A(v))g(x)\\
&-g(x)^{-1}A(v)_xdg(w)\\
&-g(x)^{-1}A([v,w])g(x).
\end{aligned} \tag{7.54}$$

We can regroup the terms two by two in order to form dA_α and some wedge; with simpler notations, we can write :

$$d(g^{-1}A_\alpha g) = -(g^{-1}dg\, g) \wedge (A_\alpha g) - (g^{-1}A) \wedge dg + (g^{-1}dAg). \tag{7.55}$$

We compute $d(g^{-1}dg)$ in the same way; the result is

$$(g^{-1}dg)(v)_x = \frac{d}{dt}\Big[g(x)^{-1}g(v_t)\Big]_{t=0} \in \mathcal{G}.$$

For v, $w \in \mathfrak{X}(M)$, we have :

$$\begin{aligned}
w\big((g^{-1}dg)(v)\big) &= \frac{d}{du}(g^{-1}dg)(v)_{w_u}\Big|_{u=0}\\
&= \frac{d}{du}\frac{d}{dt}\Big[g(w_u)^{-1}g(v_{w_u}(t))\Big]_{\substack{t=0\\u=0}}\\
&= \frac{d}{dt}\frac{d}{du}\Big[g(w_u)^{-1}g(v_t)\Big]_{\substack{u=0\\t=0}} + \frac{d}{dt}\frac{d}{du}\Big[g(x)^{-1}g(w_u(t))\Big]_{\substack{u=0\\t=0}}\\
&= d(g^{-1})(w)dg(v) + \frac{d}{du}\Big[g(x)^{-1}dg(v_{w_u})\Big]_{u=0}\\
&= -g^{-1}dg(w)g^{-1}dg(v) + g(x)^{-1}w(dg(v))
\end{aligned} \tag{7.56}$$

where w_u is a path such that $w_0' = w_x$ and $v_{w_u}(t)$ is, with respect of t, a path which gives the vector v_{w_u}. On the another hand, we have

$$(g^{-1}dg)([v,w]) = g^{-1}dg([v,w]).$$

Remark that the term $g(x)^{-1}w(dg(v))$ of $w((g^{-1}dg)(v))$ together with the same with $v \leftrightarrow w$ and $(g^{-1}dg)([v,w])$ which comes from $(g^{-1}dg)([v,w])$ will give $g(x)^{-1}(d^2g)(v,w) = 0$ when we will compute $d(g^{-1}dg)$. Finally,

$$d(g^{-1}dg) = -(g^{-1}dg\, g^{-1} \wedge dg). \tag{7.57}$$

The equations (7.55) and (7.57) allow us to write :

$$\begin{aligned}
(dA_\beta) &= d(g^{-1}A_\alpha g) + d(g^{-1}dg)\\
&= -(g^{-1}dg\, g^{-1}) \wedge (A_\alpha g) - (g^{-1}A_\alpha) \wedge dg\\
&\quad + (g^{-1}dA_\alpha g) - (g^{-1}dg\, g^{-1}) \wedge dg.
\end{aligned} \tag{7.58}$$

Notice that the term $(g^{-1}dA_\alpha g)$ corresponds to the first one in $F_\beta = g^{-1}(dA_\beta + A_\beta \wedge A_\beta)g$.

For anyone who had understood the whole computations up to here, it is clear that

$$\begin{aligned}
[A_\beta(v), A_\beta(w)] = &\frac{d}{dt}\frac{d}{du}\Big[e^{tA_\beta(v)}e^{tA_\beta(w)}\Big]_{\substack{u=0\\t=0}}\\
&- \frac{d}{dt}\frac{d}{du}\Big[e^{tA_\beta(w)}e^{tA_\beta(v)}\Big]_{\substack{u=0\\t=0}},
\end{aligned} \tag{7.59}$$

so that

$$\begin{aligned}
A_\beta \wedge A_\beta = g^{-1}A_\alpha g \wedge g^{-1}A_\alpha g &+ g^{-1}A_\alpha g \wedge g^{-1}dg\\
+ g^{-1}dg \wedge g^{-1}A_\alpha g &+ g^{-1}dg \wedge g^{-1}dg.
\end{aligned} \tag{7.60}$$

Lemma 7.13 allows us to write it under the form

$$\begin{aligned}
A_\beta \wedge A_\beta = g^{-1}A_\alpha g \wedge g^{-1}A_\alpha g &+ g^{-1}A_\alpha g \wedge g^{-1}dg\\
+ g^{-1}dg \wedge g^{-1}A_\alpha g &+ g^{-1}dg \wedge g^{-1}dg.
\end{aligned} \tag{7.61}$$

Here the term $(g^{-1}A_\alpha \wedge A_\alpha g)$ corresponds to the second one in $F_\beta = g^{-1}(dA_\beta + A_\beta \wedge A_\beta)g$. The sum of (7.58) and (7.61) is

$$F_\beta = g^{-1}F_\alpha g.$$

□

7.4.3 The electromagnetic field F

Now, we are able to interpret the field F introduced in equation (7.25). We follow [48]. From now, we use the usual Minkowski metric $g = diag(-,+,+,+)$. From the vector given by (7.27), we define a (local) potential 1-form

$$A = A_\mu dx^\mu = -\phi dt + A_x dx + A_y dy + A_z dz.$$

The field strength is $F = dA$. We easily find that

$$\begin{aligned} F = &(dt \wedge dx)(\partial_x \phi + \partial_t A_x) + \ldots \\ &+ (dx \wedge dy)(-\partial_z A_x + \partial_x A_y) + \ldots \end{aligned} \tag{7.62}$$

But the fields $\mathbf{B}$ and $\mathbf{E}$ are defined from $\mathbf{A}$ and ϕ by (7.23), so

$$\begin{aligned} F = &-E_x(dt \wedge dx) - E_y(dt \wedge dy) - E_z(dt \wedge dz) \\ &+ B_x(dy \wedge dz) + B_y(dz \wedge dx) + B_z(dx \wedge dy). \end{aligned} \tag{7.63}$$

We naturally have $dF = d^2 A = 0$. But conversely, $dF = 0$ ensures the existence of a 1-form A such that $F = dA$. If we define[3] $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = -\nabla\phi - \partial_t \mathbf{A}$, equations (7.22b) and (7.22c) are obviously satisfied. So in the connection formalism, the equations "without sources" are written by

$$dF = 0. \tag{7.64}$$

In order to write the two others, we introduce the current 1-form :

$$j = j_\mu dx^\mu = -\rho dt + j_x dx + j_y dy + j_z dz.$$

One sees that

$$\begin{aligned} \delta F := \star d \star F = &-dt(\nabla \cdot \mathbf{E}) \\ &+ dx(-\partial_t \mathbf{E}_x + (\nabla \times \mathbf{B})_x) \\ &+ dy(-\partial_t \mathbf{E}_y + (\nabla \times \mathbf{B})_y) \\ &+ dz(-\partial_t \mathbf{E}_z + (\nabla \times \mathbf{B})_z), \end{aligned} \tag{7.65}$$

so that equation $\delta F = j$ gives equations (7.22a) and (7.22d). Now, the complete set of Maxwell's equations is :

$$dF = 0 \tag{7.66a}$$
$$\delta F = j \tag{7.66b}$$

with

$$j = -\rho dt + j_x dx + j_y dy + j_z dz, \tag{7.67a}$$
$$\mathbf{B} = \nabla \times \mathbf{A} \tag{7.67b}$$
$$\mathbf{E} = -\nabla\phi - \partial_t \mathbf{A} \tag{7.67c}$$

where A is a 1-form such that $F = dA$ whose existence is given by (7.66a).

7.5 Inclusion of the Lorentz group

Up to now we had seen how to express the *gauge* invariance of a physical theory. In particle physics, a really funny field theory must be invariant under the Lorentz group; it is rather clear that, from the bundle point of view, this feature will be implemented by a Lorentz-principal bundle and some associated bundles. A spinor will be a section of an associated bundle for spin one half representation of the Lorentz group on $\mathbb{C}^4$. In order to describe non-zero spin particle interacting with an electromagnetic field (represented by a connection on a $U(1)$-principal bundle), we have to build a correct $\mathrm{SL}(2,\mathbb{C}) \times U(1)$-principal bundle. We are going to use the ideas of 6.2.1.

A **space-time** is a differentiable **pseudo-Riemannian** 4-dimensional manifold. The pseudo-Riemannian structure is a 2-form $g \in \Omega^2(M)$ for which we can find at each point $x \in M$ a basis $b = (\mathbf{b}_0, \ldots, \mathbf{b}_3)$ which fulfils

$$g_x(\mathbf{b}_i, \mathbf{b}_j) = \eta_{ij}.$$

When we use an adapted coordinates, the metric reads $g = \eta_{ij} dx^i \otimes dx^j$.

[3] *i.e.* we consider F as the main physical field while $\mathbf{E}$ and $\mathbf{B}$ are "derived" fields.

One says that M is **time orientable** if one can find a vector field $T \in \mathfrak{X}(M)$ such that $g_x(T_x, T_x) > 0$ for all $x \in M$. A **time orientation** is a choice of such a vector field. A vector $v \in T_xM$ is **future directed** if $g_x(T_x, v) > 0$.

The Lorentz group L acts on the orthogonal basis of each T_xM, but you may note that L don't act on M; it's just when the metric is flat that one can identify the whole manifold with a tangent space and consider that L is the space-times isometry group. In the case of a curved metric, the Lorentz group have to be introduced pointwise and the building of a frame bundle is natural.

Now, we are mainly interested in the frame related each other by a transformation of $L^{\uparrow}_{+}$. An arising question is to know if one can make a choice of some basis of each T_xM in such a manner that

(i) pointwise, the chosen frames are related by a transformation of $L^{\uparrow}_{+}$,

(ii) the choice is globally well defined.

The first point is trivial to fulfil from the definition of a space-time. For the second, it turns out that a good choice can be performed if and only if there exists a vector field $V \in \mathfrak{X}(M)$ such that $g_x(V_x, V_x) > 0$ for all $x \in M$. We suppose that it is the case[4].

So our first principal bundle attempt to describe the space-time symmetry is the $L^{\uparrow}_{+}$-principal bundle of orthonormal oriented frame on M :

$$\begin{tikzcd} L^{\uparrow}_{+} \arrow[r, rightsquigarrow] & L(M) \arrow[d, "p_L"] \\ & M \end{tikzcd} \tag{7.68}$$

The notion of "**relativistic invariance**" has to be understood in the sense of associated bundle to this one. The next step is to recall ourself (see subsection 6.2.1) that the physical fields doesn't transform under representation of the group $L^{\uparrow}_{+}$ but rather under representations of $\mathrm{SL}(2, \mathbb{C})$. So we build a $\mathrm{SL}(2, \mathbb{C})$-principal bundle

$$\begin{tikzcd} \mathrm{SL}(2, \mathbb{C}) \arrow[r, rightsquigarrow] & S(M) \arrow[d, "p_S"] \\ & M \end{tikzcd}$$

In order this bundle to "fit" as close as possible the bundle (7.68), we impose the existence of a map $\lambda \colon S(M) \to L(M)$ such that

(i) $p_B(\lambda(\xi)) = p_S(\xi)$ for all $\xi \in S(M)$ and

(ii) $\lambda(\xi \cdot g) = \lambda(\xi) \cdot \mathrm{Spin}(g)$ for all $g \in \mathrm{SL}(2, \mathbb{C})$.

You can recognize the definition of a **spin structure**. Notice that the existence of a spin structure on a given manifold is a non trivial issue.

Now a physical field is given by a section of the associated bundle $E = S(M) \times_\rho V$ where ρ is a representations of $\mathrm{SL}(2, \mathbb{C})$ on V. For an electron, it is $V = \mathbb{C}^4$ and $\rho = D^{(1/2,0)} \oplus D^{(0,1/2)}$. That describes a *free* electron is the sense that it doesn't interacts with a gauge field. So in order to write down the formalism in which lives a non zero spin particle, we have to build a $U(1) \times \mathrm{SL}(2, \mathbb{C})$-principal bundle. For this, we follow the procedure given in section 4.12

7.6 Interactions

7.6.1 Spin zero

The general framework is the following :

$$\begin{tikzcd} U(1) \arrow[r, rightsquigarrow] & P \arrow[d, "\pi"'] & & E = P \times_\rho V \\ & M & \mathcal{U}_\alpha \arrow[l] \arrow[ul, "\sigma_\alpha"'] \arrow[ur, "\phi"'] & \end{tikzcd}$$

a $U(1)$-principal bundle over a manifold M (as far as topological subtleties are concerned, we suppose $M = \mathbb{R}^4$) and a section ϕ of an associated bundle for a representation ρ of $U(1)$ on V. We consider M with the Lorentzian

[4] That condition is rather restrictive because we cannot, for example, find an everywhere non zero vector field on the sphere S^n with n even.

metric but, since we are intended to treat with scalar (spin zero) fields, we still don't include the Lorentz (or $SL(2,\mathbb{C})$) group in the picture. We also consider local sections $\sigma_\alpha\colon \mathcal{U}_\alpha \to P$, a connection ω on P and Ω its curvature. We define $A_\alpha = \sigma_\alpha^*\omega$.

Now we particularize ourself to the target space $V = \mathbb{C}$ on which we put the scalar product

$$\langle z_1, z_2\rangle = \frac{1}{2}(z_1\overline{z}_2 + z_2\overline{z}_1), \tag{7.69}$$

and the representation $\rho_n\colon U(1) \to GL(\mathbb{C})$,

$$\rho_n(g)z = g\cdot z = g^n z$$

where we identify $U(1)$ to the unit circle in $\mathbb{C}$ in order to compute the product. A property of the product (7.69) is to make ρ_n an isometry: for all $g \in U(1)$, $z_1, z_2 \in \mathbb{C}$,

$$\langle \rho_n(g)z_1, \rho_n(g)z_2\rangle = \langle z_1, z_2\rangle.$$

Our first aim is to write the covariant derivative of ϕ with respect to the connection ω. For this we work on the section ϕ under the form $\phi_{(\alpha)}\colon M \to V$ and we use formula (7.7) :

$$(D_X\phi)_{(\alpha)}(x) = X_x\phi_{(\alpha)} - \rho_*\big((\sigma_\alpha^*\omega)_x X_x\big)\phi_{(\alpha)}(x). \tag{7.70}$$

Let us study this formula. We know that $(\sigma_\alpha^*\omega)_x = A_\alpha(x) : T_x\mathcal{U}_\alpha \xrightarrow{\sigma} T_{\sigma_\alpha(x)}P \xrightarrow{\omega} u(1)$. Thus $A_\alpha(x)X_x$ is given by a path in $U(1)$; it is this path which is taken by ρ_*. Therefore (we forget some dependences in x)

$$\begin{aligned}
\rho_*\big(A_\alpha(x)X_x\big)\phi_{(\alpha)}(x) &= \frac{d}{dt}\Big[\rho_n\big((A_\alpha X)(t)\big)\phi_{(\alpha)}(x)\Big]_{t=0}\\
&= \frac{d}{dt}\Big[(A_\alpha X)(t)^n\Big]_{t=0}\phi_{(\alpha)}(x)\\
&= n\frac{d}{dt}\Big[(A_\alpha X)(t)\Big]_{t=0}\phi_{(\alpha)}(x)\\
&= nA_\alpha(X)\phi_{(\alpha)}(x).
\end{aligned} \tag{7.71}$$

Thus the covariant derivative is given by

$$(D_X\phi)_{(\alpha)}(x) = X_x\phi_{(\alpha)} - nA_\alpha(x)(X_x)\phi_{(\alpha)}(x). \tag{7.72}$$

One can guess an electromagnetic coupling for a particle of electric charge n. If this reveals to be physically relevant, it shows that the "electromagnetic identity card" of a particle is given by a representation of $U(1)$. This has to be seen in relation to the discussion on page 244 where the "type of particle" was closely related to representations of the Lorentz group. It is a remarkable piece of quantum field theory: the properties of a particle are encoded in representations of some symmetry groups.

Now we are going to prove that $\|D\phi\|^2$ is a gauche invariant quantity. The first step is to give a sense to this norm. We consider X_i ($i = 0,1,2,3$), an orthonormal basis of T_xM and we naturally denote $D_i = D_{X_i}$, $\partial_i = X_i$ and $A_{\alpha i} = A_\alpha(\partial_i)$. Remark that

$$A_\alpha(x)X_x = (\sigma_\alpha^*)_x X_x = \omega(d\sigma_\alpha X_x) = \omega\frac{d}{dt}\Big[\sigma_\alpha(X(t))\Big]_{t=0} \in u(1), \tag{7.73}$$

so this is given by a path in $U(1)$ which can be taken by ρ. Let $c(t)$ be this path, then

$$A_\alpha\phi_{(\alpha)}(x) = \frac{d}{dt}\Big[e^{ic(t)}\phi_{(\alpha)}(x)\Big]_{t=0},$$

so that under the conjugation, $\overline{A_\alpha\phi_{(\alpha)}(x)} = -A_\alpha\overline{\phi}_{(\alpha)}(x)$. Now our definition of $\|D\phi\|^2$ is a composition of the norm on V and the one on T_xM :

$$\|D\phi\|^2 = \eta^{ij}\langle D_i\phi_{(\alpha)}, D_j\phi_{(\alpha)}\rangle \tag{7.74}$$

Using the notation in which the upper indices are contractions with η^{ij}, we have

$$\|D\phi\|^2 = \Big((\partial_i\phi_{(\alpha)})(x) - nA_{\alpha i}\phi_{(\alpha)}(x)\Big)\Big((\partial^i\overline{\phi}_{(\alpha)})(x) + nA_\alpha^i\overline{\phi}_{(\alpha)}(x)\Big).$$

7.6.1.1 Gauge transformation law

A gauge transformation φ is given by an equivariant function $\tilde{\varphi}_\alpha \colon \mathcal{U}_\alpha \to U(1)$ which can be written under the form

$$\tilde{\varphi}_\alpha(x) = e^{i\Lambda(x)}$$

for a certain function $\Lambda \colon \mathcal{U}_\alpha \to \mathbb{R}$. From the general formula (ii) of lemma 7.3,

$$(\varphi \cdot \phi)_{(\alpha)}(x) = \rho_n(e^{-i\Lambda(x)})\phi_{(\alpha)}(x) = e^{-ni\Lambda(x)}\phi_{(\alpha)}(x). \tag{7.75}$$

The transformation of the gauche field A is given by equation (7.11). Let us see the meaning of the term $d\tilde{\varphi}$. For $v \in T_x\mathcal{U}_\alpha$,

$$(d\tilde{\varphi}_\alpha)_x v = \frac{d}{dt}\Big[\tilde{\varphi}_\alpha(v(t))\Big]_{t=0} = \frac{d}{dt}\Big[e^{i\Lambda(v(t))}\Big]_{t=0} = i\frac{d}{dt}\Big[\Lambda(v(t))\Big]_{t=0} e^{i\Lambda(v(0))} = i(d\Lambda)_x v e^{i\Lambda(x)}. \tag{7.76}$$

Thus $\tilde{\varphi}_\alpha^{-1}(x)(d\tilde{\varphi}_\alpha)_x = i(d\Lambda)_x$. Since $U(1)$ is abelian, $\tilde{\varphi}^{-1}A\tilde{\varphi} = A$. Finally,

$$(\varphi \cdot A)_\alpha(x) = A_\alpha(x) + i(d\Lambda)_x. \tag{7.77}$$

Now we are able to prove the invariance of $\|D\phi\|^2$. First,

$$(\varphi \cdot A)_{i\alpha}(x) = (\varphi \cdot A)_\alpha(\partial_i) = A_{i_\alpha}(x) + i(\partial_i\Lambda)(x); \tag{7.78}$$

second,

$$\partial_i\left(e^{-ni\Lambda(x)}\phi_{(\alpha)}(x)\right) = -ni(\partial_i\Lambda)(x)\phi_{(\alpha)}(x) + e^{-in\Lambda(x)}(\partial_i\phi_{(\alpha)})(x). \tag{7.79}$$

With these two results,

$$\partial_i(\varphi \cdot \phi)_{(\alpha)}(x) + n(\varphi \cdot A)_{\alpha i}(\varphi \cdot \phi)_{(\alpha)}(x) = e^{-in\Lambda(x)}(nA_{\alpha i}(x) + \partial_i\phi_{(\alpha)}(x)). \tag{7.80}$$

The Yang-Mills **field strength** is given by $F_{(\alpha)} = \sigma_\alpha^*\Omega$ (cf. page 165). Since $U(1)$ is abelian, $dF_{(\alpha)} = 0$, so that the second pair of Maxwell's equations is complete without any Lagrangian assumptions.

The full Yang-Mills action is written as

$$S(\omega, \phi) = \int_M \left[-\frac{1}{4}F_{(\alpha)ij}F_{(\alpha)}{}^{ij} + \frac{1}{2}\|D\phi\|^2 + \frac{1}{2}m\phi_{(\alpha)}\overline{\phi_{(\alpha)}}\right].$$

The Euler-Lagrange equations are

$$(\partial_i - inA_{\alpha i})(\partial^i - inA_\alpha^i)\phi_\alpha + m^2\phi_\alpha = 0 \tag{7.81a}$$
$$\partial_i F_{(\alpha)}{}^{ij} = 0. \tag{7.81b}$$

So the Yang-Mills Lagrangian only gives the first pair of Maxwell's equations while the second one is given by the geometric nature of fields.

As explained in [58], the topology of the physical space has deep implications on the physics of Yang-Mills equations. The absence of magnetic monopoles for example is ultimately linked to the (simple) connectedness of $\mathbb{R}^4$. When one consider the $U(1)$ Yang-Mills on a sphere, some topological charges appear and magnetic monopoles naturally arise.

7.6.2 Non zero spin formalism

The formalism for a non zero spin particle in an electromagnetic field is described in section 4.12. We consider the spinor bundle

$$\begin{array}{ccc} \mathrm{SL}(2,\mathbb{C}) & \rightsquigarrow & S(M) \\ & & \Big\downarrow{\scriptstyle p_S} \\ & & M \end{array}$$

with the spinor connection on $S(M)$, and ρ_1, a representation of $\mathrm{SL}(2,\mathbb{C})$ on V. For an electron, it is $V = \mathbb{C}^4$ and $\rho_1 = D^{(1/2,0)} \oplus D^{(0,1/2)}$, so for $g_1 \in \mathrm{SL}(2,\mathbb{C})$,

$$\rho_1(g_1)\begin{pmatrix} z_1 \\ \vdots \\ z_4 \end{pmatrix} = \begin{pmatrix} g_1 & \\ & (\overline{g_1}^t)^{-1} \end{pmatrix}\begin{pmatrix} z_1 \\ \vdots \\ z_4 \end{pmatrix}. \tag{7.82}$$

On the other hand, we consider the principal bundle

$$\begin{tikzcd} U(1) \arrow[r, rightsquigarrow] & P \arrow[d, "p_U"] \\ & M \end{tikzcd}$$

with a connection ω_2 which describes the electromagnetic field. As representation $\rho_2 \colon U(1) \to GL(\mathbb{C}^4)$ we choose the multiplication coordinate by coordinate :

$$\rho_2(g_2)\begin{pmatrix} z_1 \\ \vdots \\ z_4 \end{pmatrix} = \begin{pmatrix} g_2 z_1 \\ \vdots \\ g_2 z_4 \end{pmatrix}. \tag{7.83}$$

The physical picture of the electron is now the principal bundle

$$\begin{tikzcd} \mathrm{SL}(2,\mathbb{C}) \times U(1) \arrow[r, rightsquigarrow] & S(M) \circ P \arrow[d, "p"] \\ & M, \end{tikzcd}$$

and the field is a section of the associated bundle $(S(M) \circ P) \times_\rho \mathbb{C}^4$.

Appendix A

Complements

A.1 Alternative formalism for the quantum mechanics

We can a little reformulate the axioms of the quantum mechanics. Since we are in a Hilbert space $\mathscr{H}$ we can speak about orthogonal projections; if $\phi \in \mathcal{R}$, we can consider the projection on the space spanned by ϕ:

$$P_\phi e_k = \frac{\langle \phi, e_k \rangle}{\|\phi\|} \phi$$

where $\{e_i\}$ is a basis of $\mathscr{H}$. It is pretty clear that

$$\operatorname{Tr}(P_\phi P_\psi) = \frac{|\langle \psi, \phi \rangle|^2}{\|\psi\| \|\phi\|}. \tag{A.1}$$

If $\psi \in \mathcal{R}$ and $\phi \in \mathcal{R}'$ are unimodular, then

$$P(\mathcal{R} \to \mathcal{R}') = \operatorname{Tr}(P_\phi P_\psi), \tag{A.2}$$

so we can express the axioms in terms of projections instead of rays. For notational convenience, we put

$$\mathscr{H}_1 = \{\psi \in \mathscr{H} \text{ st } \|\psi\| = 1\}. \tag{A.3}$$

We denote by $\mathscr{S}$ the space of the projections into one dimensional subspaces of $\mathscr{H}$ (in other words $\mathscr{S}$ is the space of physical states) and for P, $Q \in \mathscr{S}$, the transition probability is $P \cdot Q = \operatorname{Tr}(PQ)$. Now a **quantum symmetry** is a map $T \colon \mathscr{S} \to \mathscr{S}'$ such that $(TP) \cdot (TQ) = P \cdot Q$.

One can prove the following :

Theorem A.1.
If $T \colon \mathscr{S} \to \mathscr{S}'$ is a quantum symmetry, then there exists an operator $U \colon \mathscr{H} \to \mathscr{H}'$ such that

(i) $P_{U\phi} = TP_\phi$,

(ii) $U(\xi + \eta) = U(\xi) + U(\eta)$,

(iii) $\langle U\xi, U\eta \rangle = \kappa(\langle \xi, \eta \rangle)$

where P_ψ is the projection onto the one dimensional space spanned by ψ and $\kappa \colon \mathbb{C} \to \mathbb{C}$ fulfils $\kappa(\lambda) = \lambda$ or $\kappa(\lambda) = \overline{\lambda}$ and

(iv) $U(\lambda\xi) = \kappa(\lambda)\xi$.

Here is why this implies Wigner's theorem as given by theorem 6.2. Let us consider some $\varphi_i \in \mathscr{H}$ such that $\|\varphi_i\| = 1$ and P_{φ_i}, the corresponding projections. Let

$$\Delta(P_1, P_2, P_3) = \langle \varphi_1, \varphi_2 \rangle \langle \varphi_2, \varphi_3 \rangle \langle \varphi_3, \varphi_1 \rangle.$$

It is clear that this expression doesn't depend on the choice of φ_i in its ray. We have

$$\begin{aligned} \Delta(TP_1; TP_2, TP_3) &= \Delta(P_{U\varphi_1}, P_{U\varphi_2}, P_{U\varphi_3}) \\ &= \langle U\varphi_1, U\varphi_2 \rangle \langle U\varphi_2, U\varphi_3 \rangle \langle U\varphi_3, U\varphi_1 \rangle \\ &= \kappa(\langle U\varphi_1, U\varphi_2 \rangle) \kappa(\langle U\varphi_2, U\varphi_3 \rangle) \kappa(\langle U\varphi_3, U\varphi_1 \rangle) \\ &= \kappa\big(\Delta(P_1, P_2, P_3)\big). \end{aligned} \tag{A.4}$$

We can see from this that the choice of $\kappa(\lambda) = \lambda$ or $\kappa(\lambda) = \overline{\lambda}$ is determined by the data of T if $\dim \mathscr{H} \geqslant 2$. In the case where $\dim \mathscr{H} = 1$, Δ is always equals to 1 and the equality (A.4) don't give any informations. In the case $\dim \mathscr{H} \geqslant 2$, we can choice φ_1 and φ_2 such that $\langle \varphi_1, \varphi_2 \rangle$ takes any value $z \in \mathbb{C}$ with $\|z\| \leqslant 1$. Taking $\varphi_3 = \varphi_1 + \varphi_2$, we find

$$\Delta(P_1, P_2, P_3) = z(1 + \overline{z})^2.$$

which is easily non real for a suitable choice of $z \in \mathbb{C}$. Let us suppose that we have an operator U which satisfies the theorem A.1. If $\kappa(\lambda) = \lambda$, then

$$U(z\psi + z'\phi) = U(z\psi) + U(z'\phi) = zU(\psi) + zU(\phi) \tag{A.5}$$

and

$$\langle U\psi, U\phi \rangle = \kappa(\langle \psi, \phi \rangle) = \langle \psi, \phi \rangle, \tag{A.6}$$

so that U is linear. If $\kappa(\lambda) = \overline{z}$, then

$$U(z\psi) = \overline{z}U\psi \tag{A.7}$$

and

$$\langle U\xi, U\eta \rangle = \kappa(\langle \xi, \eta \rangle) = \overline{\langle \xi, \eta \rangle}. \tag{A.8}$$

A.2 Statement of some results

This appendix is devoted to the statement of some results which are used in the text, but whose demonstration should be out of our purpose.

Theorem A.2 (Cayley-Hamilton)**.**
A square matrix on $\mathbb{R}$ or $\mathbb{C}$ satisfies its own characteristic equation. If $A \in \mathbb{M}_n(\mathbb{K})$, we consider the polynomial $p(\lambda) = \det(A - \lambda \mathbb{1})$ in λ. Thus $p(A) = 0$.

For a proof see Wikipedia.

Proposition A.3.
If M is a complex $n \times n$ matrix, then there exists an unitary matrix U such that U^*MU is upper triangular.

For a proof see [59].

Definition A.4.
*A **positive defined** matrix is a matrix B such that*

$$\sum_{ij} B_{ij}\overline{x_i}x_j \tag{A.9}$$

is real and positive for every complex vector x.

Proposition A.5.
A positive defined matrix is Hermitian.

Proof. We define the Hermitian matrices $M = (B + B^*)/2$ and $N = (B - B^*)/2i$, so $B = M + iN$ and

$$\bar{x}Bx = \bar{x}Mx + i\bar{x}Nx. \tag{A.10}$$

The matrices M and B being Hermitian, the numbers $\bar{x}Mx$ and $\bar{x}Nx$ are real. If $\bar{x}Bx$ has to be real, we need $\bar{x}Nx = 0$ for every x. This shows that $N = 0$, so that $B = N$. $\square$

Theorem A.6.
Let G be a Lie group and H a subgroup (with no special other structures) of G. If H is a closed subset of G then there exists an unique analytic structure on H such that H is a topological Lie subgroup of G.

This comes from [3], chapter 2, theorem 2.3.

Lemma A.7.
Let G be a connected Lie group with Lie algebra $\mathcal{G}$ and let φ be an analytic homomorphism of G into a Lie group X with Lie algebra $\mathcal{X}$. Then

(i) The kernel $\varphi^{-1}(e)$ is a topological Lie subgroup of G. Its Lie algebra is the kernel of $d\varphi_e$.

(ii) The image $\varphi(G)$ is a Lie subgroup of $\mathcal{X}$ with Lie algebra $d\varphi(\mathcal{G}) \subset \mathcal{X}$.

This comes from [3], chapter 2, lemma 5.1.

Lemma A.8.
Let G and H be two Lie group, whose Lie algebra are $\mathcal{G}$ and $\mathcal{H}$. If $\theta\colon G \to H$ is a surjective map, then we have $\mathcal{H} \simeq \mathcal{G}/Ker\, d\theta_e$.

Theorem A.9.
Let us consider $Ad\colon SU(2) \to GL(3)$, $Ad(U)X = UXU^{-1}$. We have the following properties:

(i) Ad is a linear homomorphism,

(ii) it takes his values in $\mathrm{SO}(3)$; then we can write $Ad\colon SU(2) \to \mathrm{SO}(3)$,

(iii) it is surjective,

(iv) $Ker\, Ad = \mathbb{Z}_2$,

(v) all these properties show that

$$\mathrm{SO}(3) = \frac{SU(2)}{\mathbb{Z}_2}.$$

Corollary A.10.
An useful formula:

$$\mathrm{Ad}(e^X) = e^{\mathrm{ad}\, X}.$$

Corollary A.11.
Another useful corollary of lemma 2.13 is the particular case $\phi = \boldsymbol{Ad}(e^X)$:

$$e^X e^Y e^{-X} = e^{Ad(e^Y)X}.$$

Definition A.12.
*If (a_k) is a sequence in $\mathbb{R}$, its **upper limit** is the real number*

$$\limsup_{n\to\infty} a_n = \lim_{l\to\infty} \sup\{a_k : k \geqslant l\}.$$

Lemma A.13.
If ω is a k-form (not specially a symplectic one), and ∇ a torsion free connection, one has

$$(d\omega)(X_0,\ldots,X_k) = \sum_{i=0}^{k}(-1)^i(\nabla_{X_i}\omega)(X_0,\ldots,\hat{X}_i,\ldots X_k). \tag{A.11}$$

Remark A.14.
The link between d and ∇ comes from the fact that in the left hand side of (A.11) appears some commutators $[X_i, X_j]$, but since the connection is torsion-free,

$$[X_i, X_j] = \nabla_{X_i}X_j - \nabla_{X_j}X_i$$

The main consequence of this lemma is that $\nabla\omega = 0$ implies $d\omega = 0$.

Proposition A.15.
Consider a function $f\colon X \times E \to \overline{\mathbb{R}}$ and $z_0 \in E$ such that

- *for all $z \in E$, the function $x \to (x,z)$ is integrable,*
- *for (almost) all $x \in X$, the function $z \to f(x,z)$ is continuous at z_0,*
- *there exists a function $g \geqslant 0$ such that for all $z \in E$, $|f(x,z)| \leqslant g(x)$ almost everywhere in X.*

Then the function $h\colon E \to \mathbb{R}$ defined by $h(z) = \int_X f(x,z)$ is continuous at z_0.

Bibliography

[1] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol. I.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1996. Reprint of the 1963 original, A Wiley-Interscience Publication.

[2] J. Madore. *An Introduction to noncommutative differential geometry and its physical application.* Cambridge University press, 1999.

[3] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 80 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978.

[4] J.E. Marsden. A book about symplectic geometry and mechanics divided into several *pdf* that I found on the personal website
http://www.cds.caltech.edu/~marsden.
I noticed unfortunately that the link does non more exists.

[5] P.W. Michor. *Topics in differential geometry.* 2003. These notes are from a lecture course Differentialgeometrie und Lie Gruppen.
http://www.mat.univie.ac.at/~michor/dgbook.pdf.

[6] Anthony W. Knapp. *Representation theory of semi-simple groups*, volume 36 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1986. An overview based on examples.

[7] Philipp Fahr Dierk. Enveloping algebras of finite dimensional nilpotent lie algebras. 2003.
http://www.math.uni-bielefeld.de/~philfahr/enveloping/enveloping.html, or
http://www.math.uni-bielefeld.de/~philfahr/download/enveloping.pdf.

[8] Hans Samelson. Notes on Lie algebras. 1989.
http://www.math.cornell.edu/~hatcher/Other/Samelson-LieAlg.pdf.

[9] Shlomo Sternberg. Lie algebras. 2004.
http://www.math.harvard.edu/~shlomo/index.html.

[10] J.J. Duistermaat and J.A.C Kolk. *Lie groups.* Springer, 2000.

[11] Dragan Miličić. Lectures on lie groups. 2008-2009.
http://www.math.utah.edu/~milicic/Eprints/lie.pdf.

[12] Jürgen Berndt. Lie group actions on manifolds. 2002.
http://www.mth.kcl.ac.uk/~berndt/sophia.pdf.

[13] Anthony W. Knapp. *Lie groups beyond an introduction*, volume 140 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1996.

[14] E.P. van den Ban. Lie groups. 2003.
http://www.math.uu.nl/people/ban/lecnot.html.

[15] Arthur A. Sagle and Ralph E. Walde. *Introduction to Lie groups and Lie algebras.* Academic Press, New York, 1973. Pure and Applied Mathematics, Vol. 51.

[16] F. Wisser. Classification of complex and real semi-simple Lie algebras. 2001. Some good material. Among other, root space decomposition and Cartan subalgebras.
http://www.mat.univie.ac.at/~cap/files/wisser.pdf.

[17] G. Hochschild. *The structure of Lie groups.* Holden-Day Inc., San Francisco, 1965.

[18] Jean-Pierre Serre. *Algèbres de Lie semi-simples complexes*. W. A. Benjamin, inc., New York-Amsterdam, 1966.

[19] J.F. Cornwell. *Group theory in physics, volume 2*. 1984.

[20] Thomas R. Covert. The Xartan matrix of a root system. 2005.
http://ocw.mit.edu/courses/mathematics/18-06ci-linear-algebra-communications-intensive-spring-2004/

[21] Brian G. Wybourne. *Classical groups for physicists*. 1974. Very complete discussion about computation of root and weight space; many examples on $\mathfrak{su}(3)$.

[22] Robert N Cahn. Semi-simple lie algebras and their representations. 1984. http://phyweb.lbl.gov/ rncahn/www/liealgebras/texall.pdf.

[23] James E. Humphreys. *Reflexion groups and Coxeter groups*. 1997. Some pages are available at
http://books.google.com/books?id=ODfjmOeNLMUC&hl=en.

[24] Nicolas Boulanger, Sophie de Buyl, and Francis Dolan. Semi-simple Lie algebras and representations. 2005.
http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveI/Modave2005.pdf.

[25] James E. Humphreys. *Introduction to Lie algebras and representation theory*, volume 9 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1978. Second printing, revised.

[26] D. P. Želobenko. *Compact Lie groups and their representations*. American Mathematical Society, Providence, R.I., 1973. Translated from the Russian by Israel Program for Scientific Translations, Translations of Mathematical Monographs, Vol. 40.

[27] Vicror Piercey. Verma modules. 2006.
http://math.arizona.edu/~vpiercey/VermaModules.pdf.

[28] Singh Anupam. Representation theory of Lie algebras. 2007. Workshop on Group Theory 18 Dec 2006 - 5 Jan 2007, Bangalore
http://www.isibang.ac.in/~statmath/conferences/gt/lie_algebra.pdf.

[29] New groups from old.
http://www.oup.co.uk/pdf/0-19-853548-1.pdf.

[30] Allen Hatcher. Vector bundles and K-theory. 2003.
http://www.math.cornell.edu/~hatcher/VBKT/VBpage.html.

[31] Gregory L. Naber. *Topology, geometry, and gauge fields*, volume 141 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2000. Interactions.

[32] J. Dieudonné. *Éléments d'analyse. Tome IV. Chapitres XVIII à XX*. Gauthier-Villars, Paris, 1977. Nouveau tirage, Cahiers Scientifiques, Fasc. 34.

[33] Mathew A. Johnson. The hodge decomposition theorem. Adapted from Warners book "Foundations of differential manifolds and Lie groups"
http://www.math.uiuc.edu/~mjohns51/Presentations/Hodge_Theorem.pdf.

[34] Sylvain Cappell, Dennis DeTurck, Herman Gluck, and Edward Y. Miller. Cohomology of harmonic forms on riemannian manifolds with boundary.
http://www.math.upenn.edu/~deturck/papers/har-coho-7.pdf.

[35] Vladimir G. Ivancevic and Tijana T. Ivancevic. Undergraduate lecture notes in De Rham-Hodge theory. 2008.
arXiv:0807.4991.

[36] P. Bieliavsky. Équation de champs dans l'espace de minkowski et son compactifié. Master's thesis, Université libre de Bruxelles, 1991.

[37] H. Blaine Lawson, Jr. and Marie-Louise Michelsohn. *Spin geometry*, volume 38 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1989.

[38] Joseph C. Várilly and Pawel Witkowski. *Dirac Operators and Spectral Geometry*. 2006.
http://toknotes.mimuw.edu.pl/sem3/files/Varilly_dosg.pdf.

[39] Blake Mellor. Spin$^{\mathbb{C}}$-manifolds.
http://math.berkeley.edu/~alanw/ps/mellor.ps.

[40] Thomas P. Branson and Peter B. Gilkey. Residues of the eta function for an operator of Dirac type. *J. Funct. Anal.*, 108(1):47–87, 1992.
http://www.math.uiowa.edu/ftp/branson/paper24.ps.

[41] W. Fulton and J. Harris. *Representation theory*. Springer-Verlag New York, Inc, 1991.

[42] C. Chevalley. *The Alegraic theory of spinors*. 1954. A deep discussion about Clifford algebras.

[43] Nick Brönn. Dirac operators. May 2003. A good explanation of the links between Clifford algebra and the problem of "square root" for the Laplace operator. It is done for the Euclidean metric and in Riemannian spaces. I found it on the internet but I don't remember where.

[44] S.Weinberg. *The quantum theory of fields*, volume 1. 1995.

[45] M.E. Peskin and D.V.Schroeder. *An introduction to quantum field theory*. 1995. Title says what it is.

[46] F.Schwabl. *Advanced quantum mechanics*. 1999.

[47] C. Schomblond. *Électrodynamique classique*. 2003–2004. Notes du cours d'électromagnétisme de deuxième année en physique à l'université libre de Bruxelles.
http://homepages.ulb.ac.be/~cschomb/notes.html.

[48] G.Svetlichny. Preparation for gauge theory. 1999. Almost all you need —and wish— to know about differential geometry in order to understand the gauge theories of the mathematical physics: Lie groups, fibre bundles, connections, gauge transformation, spin bundle and so on,
arXiv:math-ph/9902027.

[49] Clément Hyvrier. Étude géométrique de certaines variétés localement anti-de Sitter: les trous noirs «BTZ» massifs et sans moment angulaire. 2003. Mémoire de DEA, Université libre de Bruxelles, faculté des sciences, département de mathématiques.

[50] A. Franc. *Structures de spin et opérateurs de Dirac sur les espaces riemanniens symmétriques compacts simplement connexes*. PhD thesis, Université Libre de Bruxelles, 1988—1989.

[51] S. Sternberg. *Group theory and physics*. Cambridge University Press, Cambridge, 1994.

[52] Sami Virtanen. Quantum mechanics 3.
http://www.fyslab.hut.fi/kurssit/Tfy-44.135/.

[53] David Sénéchal. Mécanique quantique. 2000.
http://www.physique.usherbrooke.ca/senechal/doc/PHY731.pdf.

[54] Jacques Faraut. Groupes et algèbres de lie. 2001.
http://les-mathematiques.u-strasbg.fr/phorum/download.php/2,5426/GroupesdeLie.ps
This postscript file is actually a mess while a transformation into pdf *via* ps2pdf works well.

[55] Paul ES Wormer. Angular momentum theory.
http://www.theochem.ru.nl/~pwormer/angmom.pdf.

[56] Xavier Bekaert and Nicolas Boulanger. *The unitary representations of the Poincaré group in any spacetime dimension*. 2006. Lecture notes for the second edition of the Modave summer school
http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveII/Modave2006.pdf.

[57] L.Landau and E.Lifchitz. *Physique Théorique*. 1989. Théorie des champs.

[58] A. Collinucci and A. Wijns. *Topology of fiber bundles and instantons*. 2006. Lecture notes for the second edition of the Modave summer school
http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveII/Modave2006.pdf.

[59] Denis Serre. *Matrices*. Graduate Texts in Mathematics. Springer, 2010.

List of symbols

Algebra

(α, β) Inner product on the dual $\mathfrak{h}^*$ of a Cartan algebra, page 63

$\mathfrak{A}^\times$ The set of invertible elements of the algebra $\mathfrak{A}$; for example $\mathbb{R}^\times = \mathbb{R}\backslash\{0\}$, page 188

χ A representation of $\Gamma(p,q)$, page 186

Δ Basis of the roots, page 103

$\tilde{\rho}\colon (\mathbb{R}^2)^{\mathbb{C}} \to \mathrm{End}(\Lambda W)$ Spinor representation, page 194

$\tilde{\rho}\colon (\mathbb{R}^{1+3})^{\mathbb{C}} \to \mathrm{End}(\Lambda W)$ Spinor representation, page 183

$N\colon \Gamma(p,q) \to \Gamma(p,q)$ Spin norm, page 186

$\mathbb{H}$ quaternionic algebra, page 194

$\Gamma(E)$ Space of sections of the vector bundle E, page 145

γ_i Abstract definition of Dirac matrices, page 184

γ_i Explicit form of gamma matrices, page 185

$\mathrm{Ind}_{\mathfrak{A}}^{\mathrm{Cl}(V)}(E_1)$ Induced Clifford module, page 196

$\mathrm{Irr}_{\mathfrak{g}}(\mathfrak{g})$ the unique cyclic highest weight $\mathfrak{g}$-module with highest weight λ., page 134

ΛW Space of spinor representation, page 183

$\Lambda W^\pm$ Spinor space, page 194

$\mathfrak{g}^x$ An algebra derived from $\mathfrak{g}$, page 116

$\mathfrak{n}$ Restricted roots, page 110

$\Omega(M,V)$ V valued 1-forms, page 18

ω_i^j Connection form, page 161

Φ, Φ^+ Root system, page 103

$\mathrm{Ray}\,\mathscr{H}$ Rays in a Hilbert space, page 233

$\mathcal{Q}$ A subgroup of $\mathcal{G}$, page 189

A_α Gauge potentials, page 245

d_ω Exterior covariant derivative associated with the connection form ω, page 165

E_{ij} Matrix full of zero's and 1 at position ij, page 53

$F_{\mu\nu}$ Electromagnetic field strength, page 249

J_μ Electromagnetic 4-current, page 249

$U(\mathcal{A})$ Universal enveloping algebra, page 119

$V^{\mathbb{C}}$ Complexification of V, page 104

$W, \underline{W}$ Totally isotropic subspace, page 182

$W^{\mathbb{R}}$ Restriction of a complex vector spaces to $\mathbb{R}$, page 104

X^* Image of a tensor in the universal enveloping algebra, page 119

$x^\perp$ Space orthogonal to x, page 187

Differential geometry

$(\theta_\alpha)_i^j$ Matrix associated with a connection, page 158

$[\omega \wedge \eta]$ Combination of the wedge and the bracket in the case of Lie algebra-valued forms, page 146

Ad Adjoint representation, page 41

$\mathrm{Ad}(P)$ Adjoint bundle of the principal bundle P, page 157

Δ Laplace-Beltrami operator, page 175

$\mathcal{D}$ Dirac operator, page 202

$\gamma\colon \mathfrak{X}(M) \to \operatorname{End}\Gamma(\mathcal{S})$ A key ingredient for Dirac operator, page 201

$\tilde{\nabla}\colon \mathfrak{X}(M) \times \Gamma(\mathcal{S}) \to \Gamma(\mathcal{S})$ Covariant derivative for the spinor connection, page 200

$\Omega(M, E)$ the set of E-valued differential forms, page 145

Functional analysis

Δf Laplace operator, page 160

∇f Gradient of the function f, page 160

$\nabla \cdot X$ divergence of the vector field X, page 160

Lie groups and algebras

$\operatorname{Aut}\mathfrak{a}$ Group of automorphism of $\mathfrak{a}$, page 39

$\mathrm{Cl}(2)$ Clifford algebra of $\mathbb{R}^2$, page 194

$\mathrm{Cl}(p, q)$ Clifford algebra of $\mathbb{R}^{1,3}$, page 179

$\mathrm{Cl}(p, q)^\pm$ Grading of Clifford algebra, page 186

$\Gamma(p, q)$ Clifford group, page 186

$\mathrm{Int}(\mathfrak{a})$ Adjoint group of $\mathfrak{a}$, page 39

$\mathfrak{Lie}(\Gamma(p, q)^+)$ Algèbre de $\Gamma(p, q)^+$, page 193

$\mathcal{Z}(\mathfrak{h})$ the centralizer of $\mathfrak{h}$, page 59

$\partial\mathfrak{a}$ The Lie algebra of $\mathrm{Aut}(\mathfrak{a})$, page 39

$\mathfrak{gl}(\mathfrak{a})$ space of endomorphism with usual bracket, page 39

$\mathrm{Spin}(p, q)$ Spin group of $\mathbb{R}^{1,3}$, page 186

$\mathfrak{spin}(p, q)$ Lie algebra of the group $\mathrm{Spin}(p, q)$, page 192

$\mathrm{Spin}(V)$ The spin group, page 192

$\mathrm{Spin}^c(V)$ A group related to Spin, page 191

$A \triangleleft B$ A is a normal subgroup of B, page 137

$GL(\mathfrak{a})$ The group of nonsingular endomorphism of $\mathfrak{a}$, page 39

(α, β) inner product on the dual $\mathfrak{h}^*$., page 63

$l(w)$ length in the Weyl group, page 102

t_α a basis of $\mathfrak{h}$, page 63

Index

Bibliography

[1] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol. I.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1996. Reprint of the 1963 original, A Wiley-Interscience Publication.

[2] J. Madore. *An Introduction to noncommutative differential geometry and its physical application.* Cambridge University press, 1999.

[3] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*, volume 80 of *Pure and Applied Mathematics.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978.

[4] J.E. Marsden. A book about symplectic geometry and mechanics divided into several *pdf* that I found on the personal website
http://www.cds.caltech.edu/~marsden.
I noticed unfortunately that the link does non more exists.

[5] P.W. Michor. *Topics in differential geometry.* 2003. These notes are from a lecture course Differentialgeometrie und Lie Gruppen.
http://www.mat.univie.ac.at/~michor/dgbook.pdf.

[6] Anthony W. Knapp. *Representation theory of semi-simple groups*, volume 36 of *Princeton Mathematical Series.* Princeton University Press, Princeton, NJ, 1986. An overview based on examples.

[7] Philipp Fahr Dierk. Enveloping algebras of finite dimensional nilpotent lie algebras. 2003.
http://www.math.uni-bielefeld.de/~philfahr/enveloping/enveloping.html, or
http://www.math.uni-bielefeld.de/~philfahr/download/enveloping.pdf.

[8] Hans Samelson. Notes on Lie algebras. 1989.
http://www.math.cornell.edu/~hatcher/Other/Samelson-LieAlg.pdf.

[9] Shlomo Sternberg. Lie algebras. 2004.
http://www.math.harvard.edu/~shlomo/index.html.

[10] J.J. Duistermaat and J.A.C Kolk. *Lie groups.* Springer, 2000.

[11] Dragan Miličić. Lectures on lie groups. 2008-2009.
http://www.math.utah.edu/~milicic/Eprints/lie.pdf.

[12] Jürgen Berndt. Lie group actions on manifolds. 2002.
http://www.mth.kcl.ac.uk/~berndt/sophia.pdf.

[13] Anthony W. Knapp. *Lie groups beyond an introduction*, volume 140 of *Progress in Mathematics.* Birkhäuser Boston Inc., Boston, MA, 1996.

[14] E.P. van den Ban. Lie groups. 2003.
http://www.math.uu.nl/people/ban/lecnot.html.

[15] Arthur A. Sagle and Ralph E. Walde. *Introduction to Lie groups and Lie algebras.* Academic Press, New York, 1973. Pure and Applied Mathematics, Vol. 51.

[16] F. Wisser. Classification of complex and real semi-simple Lie algebras. 2001. Some good material. Among other, root space decomposition and Cartan subalgebras.
http://www.mat.univie.ac.at/~cap/files/wisser.pdf.

[17] G. Hochschild. *The structure of Lie groups.* Holden-Day Inc., San Francisco, 1965.

[18] Jean-Pierre Serre. *Algèbres de Lie semi-simples complexes*. W. A. Benjamin, inc., New York-Amsterdam, 1966.

[19] J.F. Cornwell. *Group theory in physics, volume 2*. 1984.

[20] Thomas R. Covert. The Xartan matrix of a root system. 2005.
http://ocw.mit.edu/courses/mathematics/18-06ci-linear-algebra-communications-intensive-spring-2004/

[21] Brian G. Wybourne. *Classical groups for physicists*. 1974. Very complete discussion about computation of root and weight space; many examples on $\mathfrak{su}(3)$.

[22] Robert N Cahn. Semi-simple lie algebras and their representations. 1984. http://phyweb.lbl.gov/ rncahn/www/liealgebras/texall.pdf.

[23] James E. Humphreys. *Reflexion groups and Coxeter groups*. 1997. Some pages are available at
http://books.google.com/books?id=ODfjmOeNLMUC&hl=en.

[24] Nicolas Boulanger, Sophie de Buyl, and Francis Dolan. Semi-simple Lie algebras and representations. 2005.
http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveI/Modave2005.pdf.

[25] James E. Humphreys. *Introduction to Lie algebras and representation theory*, volume 9 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1978. Second printing, revised.

[26] D. P. Želobenko. *Compact Lie groups and their representations*. American Mathematical Society, Providence, R.I., 1973. Translated from the Russian by Israel Program for Scientific Translations, Translations of Mathematical Monographs, Vol. 40.

[27] Vicror Piercey. Verma modules. 2006.
http://math.arizona.edu/~vpiercey/VermaModules.pdf.

[28] Singh Anupam. Representation theory of Lie algebras. 2007. Workshop on Group Theory 18 Dec 2006 - 5 Jan 2007, Bangalore
http://www.isibang.ac.in/~statmath/conferences/gt/lie_algebra.pdf.

[29] New groups from old.
http://www.oup.co.uk/pdf/0-19-853548-1.pdf.

[30] Allen Hatcher. Vector bundles and K-theory. 2003.
http://www.math.cornell.edu/~hatcher/VBKT/VBpage.html.

[31] Gregory L. Naber. *Topology, geometry, and gauge fields*, volume 141 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 2000. Interactions.

[32] J. Dieudonné. *Éléments d'analyse. Tome IV. Chapitres XVIII à XX*. Gauthier-Villars, Paris, 1977. Nouveau tirage, Cahiers Scientifiques, Fasc. 34.

[33] Mathew A. Johnson. The hodge decomposition theorem. Adapted from Warners book "Foundations of differential manifolds and Lie groups"
http://www.math.uiuc.edu/~mjohns51/Presentations/Hodge_Theorem.pdf.

[34] Sylvain Cappell, Dennis DeTurck, Herman Gluck, and Edward Y. Miller. Cohomology of harmonic forms on riemannian manifolds with boundary.
http://www.math.upenn.edu/~deturck/papers/har-coho-7.pdf.

[35] Vladimir G. Ivancevic and Tijana T. Ivancevic. Undergraduate lecture notes in De Rham-Hodge theory. 2008.
arXiv:0807.4991.

[36] P. Bieliavsky. Équation de champs dans l'espace de minkowski et son compactifié. Master's thesis, Université libre de Bruxelles, 1991.

[37] H. Blaine Lawson, Jr. and Marie-Louise Michelsohn. *Spin geometry*, volume 38 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1989.

[38] Joseph C. Várilly and Pawel Witkowski. *Dirac Operators and Spectral Geometry*. 2006.
http://toknotes.mimuw.edu.pl/sem3/files/Varilly_dosg.pdf.

[39] Blake Mellor. Spin$^{\mathbb{C}}$-manifolds.
http://math.berkeley.edu/~alanw/ps/mellor.ps.

[40] Thomas P. Branson and Peter B. Gilkey. Residues of the eta function for an operator of Dirac type. *J. Funct. Anal.*, 108(1):47–87, 1992.
http://www.math.uiowa.edu/ftp/branson/paper24.ps.

[41] W. Fulton and J. Harris. *Representation theory.* Springer-Verlag New York, Inc, 1991.

[42] C. Chevalley. *The Alegraic theory of spinors.* 1954. A deep discussion about Clifford algebras.

[43] Nick Brönn. Dirac operators. May 2003. A good explanation of the links between Clifford algebra and the problem of "square root" for the Laplace operator. It is done for the Euclidean metric and in Riemannian spaces. I found it on the internet but I don't remember where.

[44] S.Weinberg. *The quantum theory of fields*, volume 1. 1995.

[45] M.E. Peskin and D.V.Schroeder. *An introduction to quantum field theory.* 1995. Title says what it is.

[46] F.Schwabl. *Advanced quantum mechanics.* 1999.

[47] C. Schomblond. *Électrodynamique classique.* 2003–2004. Notes du cours d'électromagnétisme de deuxième année en physique à l'université libre de Bruxelles.
http://homepages.ulb.ac.be/~cschomb/notes.html.

[48] G.Svetlichny. Preparation for gauge theory. 1999. Almost all you need —and wish— to know about differential geometry in order to understand the gauge theories of the mathematical physics: Lie groups, fibre bundles, connections, gauge transformation, spin bundle and so on,
arXiv:math-ph/9902027.

[49] Clément Hyvrier. Étude géométrique de certaines variétés localement anti-de Sitter: les trous noirs «BTZ» massifs et sans moment angulaire. 2003. Mémoire de DEA, Université libre de Bruxelles, faculté des sciences, département de mathématiques.

[50] A. Franc. *Structures de spin et opérateurs de Dirac sur les espaces riemanniens symmétriques compacts simplement connexes.* PhD thesis, Université Libre de Bruxelles, 1988—1989.

[51] S. Sternberg. *Group theory and physics.* Cambridge University Press, Cambridge, 1994.

[52] Sami Virtanen. Quantum mechanics 3.
http://www.fyslab.hut.fi/kurssit/Tfy-44.135/.

[53] David Sénéchal. Mécanique quantique. 2000.
http://www.physique.usherbrooke.ca/senechal/doc/PHY731.pdf.

[54] Jacques Faraut. Groupes et algèbres de lie. 2001.
http://les-mathematiques.u-strasbg.fr/phorum/download.php/2,5426/GroupesdeLie.ps
This postscript file is actually a mess while a transformation into pdf *via* ps2pdf works well.

[55] Paul ES Wormer. Angular momentum theory.
http://www.theochem.ru.nl/~pwormer/angmom.pdf.

[56] Xavier Bekaert and Nicolas Boulanger. *The unitary representations of the Poincaré group in any spacetime dimension.* 2006. Lecture notes for the second edition of the Modave summer school
http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveII/Modave2006.pdf.

[57] L.Landau and E.Lifchitz. *Physique Théorique.* 1989. Théorie des champs.

[58] A. Collinucci and A. Wijns. *Topology of fiber bundles and instantons.* 2006. Lecture notes for the second edition of the Modave summer school
http://www.ulb.ac.be/sciences/ptm/pmif/Rencontres/ModaveII/Modave2006.pdf.

[59] Denis Serre. *Matrices.* Graduate Texts in Mathematics. Springer, 2010.

List of symbols

Algebra

$W, \underline{W}$ Totally isotropic subspace, page 182

$W^{\mathbb{R}}$ Restriction of a complex vector spaces to $\mathbb{R}$, page 104

X^* Image of a tensor in the universal enveloping algebra, page 119

$x^\perp$ Space orthogonal to x, page 187

Differential geometry

$(\theta_\alpha)_i^j$ Matrix associated with a connection, page 158

$[\omega \wedge \eta]$ Combination of the wedge and the bracket in the case of Lie algebra-valued forms, page 146

Ad Adjoint representation, page 41

$\mathrm{Ad}(P)$ Adjoint bundle of the principal bundle P, page 157

Δ Laplace-Beltrami operator, page 175

$\mathcal{D}$ Dirac operator, page 202

$\gamma\colon \mathfrak{X}(M) \to \operatorname{End}\Gamma(\mathcal{S})$ A key ingredient for Dirac operator, page 201

$\tilde{\nabla}\colon \mathfrak{X}(M) \times \Gamma(\mathcal{S}) \to \Gamma(\mathcal{S})$ Covariant derivative for the spinor connection, page 200

$\Omega(M, E)$ the set of E-valued differential forms, page 145

Functional analysis

Δf Laplace operator, page 160

∇f Gradient of the function f, page 160

$\nabla \cdot X$ divergence of the vector field X, page 160

Lie groups and algebras

$\operatorname{Aut}\mathfrak{a}$ Group of automorphism of $\mathfrak{a}$, page 39

$\mathrm{Cl}(2)$ Clifford algebra of $\mathbb{R}^2$, page 194

$\mathrm{Cl}(p,q)$ Clifford algebra of $\mathbb{R}^{1,3}$, page 179

$\mathrm{Cl}(p,q)^\pm$ Grading of Clifford algebra, page 186

$\Gamma(p,q)$ Clifford group, page 186

$\mathrm{Int}(\mathfrak{a})$ Adjoint group of $\mathfrak{a}$, page 39

$\mathfrak{Lie}(\Gamma(p,q)^+)$ Algèbre de $\Gamma(p,q)^+$, page 193

$\mathcal{Z}(\mathfrak{h})$ the centralizer of $\mathfrak{h}$, page 59

$\partial\mathfrak{a}$ The Lie algebra of $\mathrm{Aut}(\mathfrak{a})$, page 39

$\mathfrak{gl}(\mathfrak{a})$ space of endomorphism with usual bracket, page 39

$\mathrm{Spin}(p,q)$ Spin group of $\mathbb{R}^{1,3}$, page 186

$\mathfrak{spin}(p,q)$ Lie algebra of the group $\mathrm{Spin}(p,q)$, page 192

$\mathrm{Spin}(V)$ The spin group, page 192

$\mathrm{Spin}^c(V)$ A group related to Spin, page 191

$A \lhd B$ A is a normal subgroup of B, page 137

$GL(\mathfrak{a})$ The group of nonsingular endomorphism of $\mathfrak{a}$, page 39

(α, β) inner product on the dual $\mathfrak{h}^*$., page 63

$l(w)$ length in the Weyl group, page 102

t_α a basis of $\mathfrak{h}$, page 63

Index